

UNITED INSTITUTE OF INFORMATICS PROBLEMS
OF THE NATIONAL ACADEMY OF SCIENCES OF BELARUS

Pattern
Recognition
and
Information
Processing

PRIP'2014

PROCEEDINGS OF THE 12th
INTERNATIONAL CONFERENCE
28–30 May 2014, Minsk, Belarus

Minsk
UIIP NASB
2014

Pattern Recognition and Information Processing (PRIP'2014) : Proceedings of the 12th International Conference (28–30 May 2014, Minsk, Belarus). – Minsk : UIIP NASB, 2014. – 340 p. – ISBN 978-985-6744-81-8.

This Proceedings book contains papers accepted for publication and presentation at the 12th International Conference “Pattern Recognition and Information Processing” that will be held in May 28–30, 2014, Minsk, Belarus.

Proceedings are prepared for publication by PRIP'2014 Program Committee and Belarusian Association for Image Analysis and Recognition. They could be useful for students and researchers working in the following areas: Pattern recognition and Image analysis, Knowledge engineering, Big data, Decision support systems, Fuzzy mathematics and Fuzzy systems, as well as Applications of pattern recognition and image analysis methods.

Papers are published in the form presented by authors who are fully responsible for correctness of the reported results and technical quality.

Сборник включает доклады 12-й Международной конференции “Распознавание образов и обработка информации”, которая состоится 28–30 мая 2014 г. в Минске (Беларусь). Сборник подготовлен к выпуску программным комитетом конференции и Белорусской ассоциацией по анализу и распознаванию изображений.

Доклады конференции будут полезны студентам и специалистам, работающим в следующих областях: распознавание образов и анализ изображений, обработка и представление знаний, “большие” данные, системы поддержки принятия решений, основанные на знаниях, нечеткая математика и системы, а также практические приложения методов распознавания образов и анализа изображений.

Статьи печатаются в виде, представленном авторами, которые несут всю ответственность за достоверность приводимых научных результатов.

Editors:

Alexander Tuzikov,
Vassili Kovalev

ISBN 978-985-6744-81-8

© United Institute of Informatics
Problems of the National Academy
of Sciences of Belarus, 2014

Organized by

United Institute of Informatics Problems of the National Academy
of Sciences of Belarus

In cooperation with

- National Academy of Sciences of Belarus
- Belarusian State University
- Belarusian State University of Informatics and Radioelectronics
- Belarusian Association for Image Analysis and Recognition
- EPAM Systems



20 Years of Excellence in Software Engineering

Welcome Words from the PRIP'2014 Chairman



On behalf of the Organizers it is my great pleasure to welcome you at the 12th International Conference “Pattern Recognition and Information Processing” (PRIP). PRIP Conference is the main biennial event of the Belarusian Association for Image Analysis and Recognition. The Conference is organized in cooperation with other major ICT research and development establishments of Belarus.

It is my pleasure to inform you that papers from 10 different countries have been submitted to PRIP'2014. All the submitted papers have been reviewed by the Program Committee members together with the external referees. As a result we have selected 66 papers for including into the PRIP'2014 Scientific Program.

The Conference can not take place without the help of many people participating in the Conference organization. I would like to express my sincere gratitude to all members of the Program Committee and referees who carefully reviewed and selected the papers, to the Organizing Committee who has been working hard for the organization of the Conference and preparation of PRIP'2014 Proceedings book, as well as to all Authors for submitting their papers.

Many thanks to everyone who took part in preparation of the Conference. I am happy to welcome you to Minsk and hope that you will have good impressions and memories of PRIP'2014 when you will open this book.

Alexander Tuzikov

PRIP'2014 Chairman
Minsk, Belarus, May 2014

PRIP'2014 COMMITTEES

Chairman of the Conference

Alexander Tuzikov, Belarus

Vice-Chairman of the Conference

Sergey Ablameyko, Belarus

Program Committee Chairman

Vassili Kovalev, Belarus

Local Organizing Committee

Viacheslau Arkhipov (Chairman)

Alexander Dmitruk

Vitali Liauchuk

Alexander Volkovich

Viktor Liakhouski

Matvey Sprindzhuk

Larisa Murashko

THE KNOWLEDGE EXTRACTION ANALYTICAL PLATFORM FOR EMBEDDED INDUSTRIAL APPLICATIONS

V. Abaturov, Yu. Dorogov
Saint Petersburg Electrotechnical University "LETI", Russia
e-mail: vasilianich@yandex.ru

The knowledge extraction analytical platform for embedded industrial applications based on DBMS PostgreSQL is considered. The architecture of analytical platform that satisfies requirements of SQL/MM and PMML is proposed. The unified management interface for analytical platform is described. The knowledge extraction phases (training phase, testing phase and application phase) are considered. The logical data model of analytical platform infrastructure is shown. The advantages of analytical platform are presented.

Introduction

Modern process control systems require data processing of large collection of information. This problem has become critical in analytical data processing fields such as data mining, artificial intelligence, decision support system (DSS), machine vision (MV), multimedia technology, etc. Currently, the analytical system market is growing exponentially. Many foreign companies such as IBM Cognos, MicroStrategy, Oracle, SAS, Microsoft, as well as Russian company BaseGroup Labs, "Prognoz" [1] work in the analytical system market.

Integration of analytical data processing, algorithms for knowledge extraction, metadata management and results visualization on the unified analytical platform software is the current trend of analytic system evolution. The technology development of analytical platform for industrial embedded applications requires solutions of several fundamental issues:

- development of architecture;
- development of system interfaces;
- providing service capabilities;
- providing security and reliability;
- providing high performance.

Unmanned technology is becoming more popular with the introduction of analytical computations in industrial systems. Other computation processes use analytical services in those technologies. Wherein, the using of universal interaction between different computation processes and analytical system is becoming more important. Transmission is supported not only at the data level, but also at the level of data processing models in system for industry. Therefore, standardization of data processing models representation is also important.

1. Standardization

Existing data mining standards affect main aspects of analytical systems for knowledge extraction building. There are three aspects that can be distinguished. First is the unification of interfaces, by which any application can access the functionality of the analytical platform. This aspect considers the object programming languages (CWM Data Mining, JDM) and SQL extensions (SQL/MM, OLE DB for Data Mining), which allow you to access Data Mining tools embedded into a relational database. Second is developing the unified agreement for storage and transmission data mining model. PMML language (Predicted Model Markup Language) is the solution of this problem. And the last one is the recommendation about processes of knowledge extraction. This aspect is considered in CRISP-DM standard (Cross Industry Standard Process for Data Mining).

The analysis shows that SQL/MM [2] and PMML [3] are the most compatible pair of standards, among many other. SQL / MM standard is sufficient for the development of the analytical platform architecture embedded into a relational database. PMML standard completely covers the presentation of algorithms and models. Besides, the last version of this standard allows using PMML models for information description.

2. The architecture of analytical platform

The analytical platform for industrial embedded applications bases on ORDBSM (Object-Relational Database Management System) PostgreSQL [4] is considered in this article. The architecture of analytical platform is shown on fig. 1.

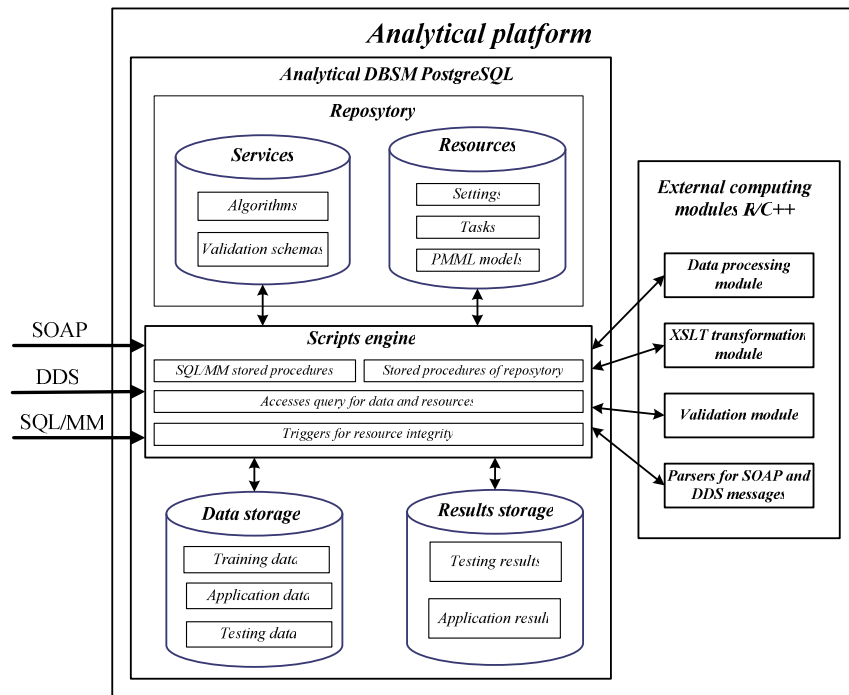


Fig. 1. The architecture of analytical platform

The analytical platform consists of the analytical DBSM and external support computing modules. The interaction between analytical platform and applications is implemented using three independent interfaces:

- SQL/MM – SQL language extension for control of knowledge extraction process;
- SOAP – communication protocol [5] between the web applications;
- DDS – distributed service for real-time systems standard (OMG) [6].

PostgreSQL performs as an analytic DBMS. PostgreSQL allows you to create a new custom data types and stored procedures to perform knowledge extraction in accordance with the standard SQL/MM.

External computing modules, in fact, represent a DLL (Dynamic Link Library) written in an object-oriented language C++ and R language (functional programming language for statistical computing). The external computing modules include: data processing module; XSLT transformation module; validation module and parsers for SOAP and DDS messages – library that is used for forming SQL/MM scripts from SOAP and DDS messages and for forming reply messages containing the result.

The interaction between analytical database and external computing modules (plug-ins) is organized by PostgreSQL standard interface for stored procedures written in C++ language and by additional PL/R module allowed to develop a stored procedure written in R language.

Analytical database contains storages for the analyzed data (training data, testing data and application data) and the results of using data mining models (test results and application results).

Centralized storage of analytical infrastructure is important aspect of analytical platform for industrial application. System storage of analytical platform infrastructure will be called a repository. The repository is the storage of analytical services and resources. Services storage contains a description of all available algorithms and all validation schemas that are necessary for processing SQL/MM procedures. Also validation schemas are used for triggers that provide the analytical resources integrity.

Resources storage contains the results of the various phases and stages of knowledge extraction. Resources include: settings – XML description of the available computational options; tasks – XML description containing all the necessary information to start the stage of knowledge retrieval and PMML models – XML description of the finished data mining models.

PMML standard is important part of analytical platform architecture. This standard applies not only for the representation and storage of data mining models. PMML standard is used to develop a message interface for transmitting information through the remote access (SOAP and DDS).

Data mining processes execution occurs on the scripts engine. The scripts engine is based on PostgreSQL request handler (SQL handler). The script engine allows not only to call the analytical and repository procedures, but also provides access to data and resources storages. SQL/MM standard (SQL query language extension) provides the unified access. In addition, the scripts engine is equipped with the resources integrity control tools. The tools are implemented by PostgreSQL trigger functions.

Exceptions are raised when the analytical platform is used incorrectly. Exceptions interrupt the scripts and retrieve the analytical platform to a previous state. Codes and messages of exceptions allow identifying the cause of incorrect use of the script. The SQL/MM script example for metadata defining is shown on fig. 2.

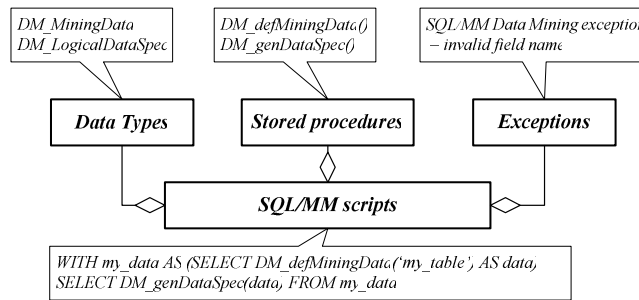


Fig. 2. The SQL/MM script example

It can be seen that the main functional elements of the SQL/MM scripts are data types, stored procedures and exceptions.

Data types are the structured units of information. Stored procedures are used to perform operations on different data types. Exceptions are used to present the information about scripts running mistakes. The aggregate of these functional elements allows creating a complete set of SQL/MM scripts for knowledge extraction running. There are stored procedures `DM_defMiningData()` and `DM_getLogicalDataSpec()`, data types `DM_MiningData` and `DM_LogicalDataSpec`, exception “SQL/MM Data Mining exception – invalid table name” is used in the fig. 2 script. “SQL/MM Data Mining exception – invalid table name” exception is generated in case of an incorrect name of the data table.

3. The infrastructure of analytical platform

The SQL/MM standard determines the common concept of knowledge extraction processes performed by specific algorithms, but does not define the analytical platform architecture generally. Nonetheless, it is contemplated that tasks, models and algorithms settings should be stored in special tables of analytical database. The set of system tables for storage the results of all knowledge extraction forms the analytical infrastructure. Analytical infrastructure should provide: access and manage of analytics platform resources (tasks, computational models and settings); specification of available services (description of the algorithms); description of available algorithms settings; description of available knowledge models; description of available SQL/MM function interfaces; integrity of analytical resources and services.

The repository is a system storage supporting the infrastructure of analytical platform. The purpose of the repository is providing the programming interface to the analytical resources and services. The logical data model of repository is shown of fig. 3. The repository consists of resources storage and services storage.

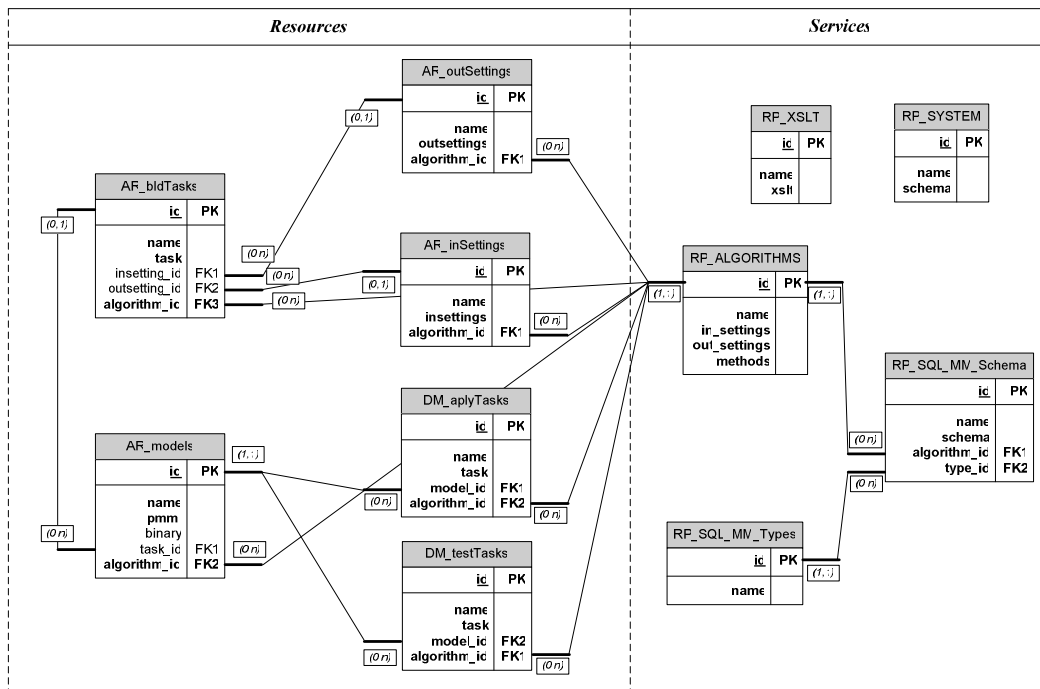


Fig. 3. The logical data model of repository

The services storage is a set of tables designed for knowledge extraction algorithms description, description of different SQL/MM types by XML schemas, storage of base validation schemas and system XSLT transformations. In general, the services storage improves the quality and efficiency of the analytical platform. The services storage simplifies the using of analytical platform and makes clear the implementation of knowledge extraction process. The names of services storage table begin with the prefix “RP_”. Fig. 3 shows that the analytical resources consist of five tables: table of knowledge extraction algorithms (RP_ALGORITHMS); table of system schemas (RP_SYSTEM); table of system transformation (RP_XSLT); table of SQL/MM types (RP_SQL_MM_Types); table of SQL/MM schemas (RP_SQL_MM_Schema).

The RP_ALGORITHMS table is intended for description of knowledge extraction algorithms. This table contains the identification fields (id, name), providing the uniqueness of algorithm, settings schemas and description of the algorithm methods. The RP_ALGORITHMS table provides preservation the resources only to those algorithms which are defined therein.

The RP_SYSTEM table designed to store basic schemas. This table contains the following basic schemas: PMML schema (*pmmml-4-1*) designed to validate the knowledge models and used as a base to the other system schemas; base schema of PMML models settings representation (*pmm_task*) used for models settings validation; schema of SQL/MM methods (*SQL/MM*) used for validation of SQL/MM methods description; schema of DDS and SOAP messages (*DDS_Message*, *SOAP_Message*) used for validation of transmitted DDS and SOAP messages.

The RP_XSLT table is intended to store system XSLT transformations used to implement SQL/MM functional.

The RP_SQL_MM_Types table is intended to store the names of abstract SQL/MM types, which include models (*Model*), training tasks (*BldTask*), testing tasks (*TestTask*) and application task (*AplyTask*). Records in this table provide uniqueness of schemes describing the internal structure of the SQL/MM types.

The RP_SQL_MM_Schema table is intended for direct storage of different SQL/MM types schemas. The combination of unique algorithms identifiers in RP_ALGORITHMS and type identifiers in RP_SQL_MM_Types guarantees, for example, that the services storage will be one and only one training task schema for the classification algorithm, of course, if the only one classification algorithm is announced in the services storage. Otherwise, the validation SQL/MM types schemas must be define for each classification.

The resources storage is a set of related tables used for storage of results on each stages of knowledge extraction process (models settings, algorithms settings, tasks, PMML models). The names resources storage tables with the prefix “AR_”. Fig. 3 shows that the resources storage consists of six

tables: table of knowledge models settings (AR_outSettings); table of knowledge extraction algorithm (AR_inSettings); table of training tasks (AR_bldTask); table of testing tasks (AR_testTask); table of application tasks (AR_aplyTask); table of knowledge models (AR_model).

The AR_outSettings table is intended to store the settings of output PMML models. The AR_inSettings table is intended to store the settings of knowledge extraction algorithms. AR_bldTask, AR_testTask and AR_aplyTask tables are intended for storage of training, testing and application task respectively. The AR_model table is designed for storing PMML models. There are not only PMML representations but although binary representations that are stored in AR_model.

4. The technical characteristics of analytical platform

Since PostgreSQL is chosen like an analytical database, the analytical platform implemented in accordance with the proposed architecture will have the following performance characteristics:

- processing of large-scale dissimilar data arrays (up to 32 TB);
- extensible and scalable of analytics;
- multiplatform;
- integrity control;
- interfaces to high-level programming languages (C, C++, R, Java);
- triggers and rules for process control;
- rights management system access and authorization;
- parallel processing of user sessions;
- traffic encryption.

Conclusion

The advantages of the analytical platform embedded into the database are the high flexibility of the knowledge extraction algorithms, as well as simple way of their expansion and scaling. The proposed architecture of the analytical platform allows to perform analytical process and knowledge extraction in different areas of industrial applications more effectively.

Implementation of the analytical platform within the database allows using scripts of data processing for the construction of cascade algorithms of any complexity. Also embedded database service provides the reliability, security, scalability, and extensibility of analytical platform that is essential for industrial use. The proposed architecture extends the SQL/MM standard in terms of analytical platform infrastructure organization. The standard assumes the using of object-oriented databases. Not all databases satisfy the requirements of SQL/MM standard fully, so the implementation of a standard dialect can be expected only. PostgreSQL is not strictly object-oriented database, but its ability adequately covers the requirements of the SQL/MM standard.

References

1. Data Mining Community. Top Resource [Electronic resource]. – Mode of access : <http://www.kdnuggets.com>. – Date of access : 16.04.2014.
2. ISO/IEC 13249-6-2006, SQL/MM Part 6.
3. Version 4.1, 2012, Data Mining Group (DMG) [Electronic resource]. – Mode of access : <http://www.dmg.org>. – Date of access : 16.04.2014.
4. PostgreSQL [Electronic resource]. – Mode of access : <http://www.postgresql.org>. – Date of access : 16.04.2014.
5. Simple Object Access Protocol, SOAP 1.2 Messaging Framework (Second Edition) [Electronic resource]. – Mode of access : <http://www.w3.org/TR/soap12-part1>. – Date of access : 16.04.2014.
6. OpenDDS Version 3.0 Supported by Object Computing, Inc. (OCI) [Electronic resource]. – Mode of access : <http://www.opendds.org/>, <http://www.ocweb.com>. – Date of access : 16.04.2014.

SPATIO-TEMPORAL CLUSTER ANALYSIS OF DISEASE

M. Abramovich, M. Mitskevich
Research Institute for Applied Problems of Mathematics and Informatics,
Minsk, Belarus
e-mail: abramovichms@bsu.by

We consider the global clusterization test and scan statistic method for studying geographical distribution of a disease. The methods were applied for childhood thyroid carcinoma cases of the Republic of Belarus.

Introduction

The methods of spatio-temporal cluster analysis can be classified as local and global [1–3]. The global tests determine a possibility of a presence of a cluster structure in total on the territory under consideration [1]. Global tests are designed to provide a single statistic that can be used to assess the degree of a map pattern deviation from null hypothesis of spatial randomness. Global tests do not provide additional information about the size and location of clusters.

By using local clusterization algorithms it is possible to find the places and sizes of clusters [2, 3]. Local methods of clustering are aimed at searching the data and uncovering the size and location of any possible clusters. For construction of spatial clusters the so-called “spherical windows” of the variable sizes are used to determine the potential territory of the cluster, so it seems to be impossible to identify the cluster of an arbitrary form by this approach. For these cases the clusters that are formed as the result often include the territories with the incidence risk that is not high. That is why algorithm of the flexibly shaped cluster construction for the cases will be considered. A flexible scan statistic imposes an irregularly shaped window on each district by connecting its adjacent districts.

If data has at least one outlier, then spatial scan statistic often determines the cluster that includes only this outlier. In these cases it can be recommended to use robust versions of statistical procedures to perform the cluster analysis adequately and correctly. We consider the robust version of the spatial scan statistic.

1. Global clusterization test

Let us suppose that the whole territory under study is divided into m districts, and for every district $c_i, i = 1, \dots, m$ is the number of cases, $n_i, i = 1, \dots, m$, is the population at risk. Denote $C = \sum_{i=1}^m c_i$ as the total number of cases, $N = \sum_{i=1}^m n_i$ as the total population at risk, $\xi_i, i = 1, \dots, m$ – random variable for the number of cases in the i -th district, $\mu_i, i = 1, \dots, m$ – expected number of cases in the i -th district. The location of each district is defined by the pair of the geographic coordinates of its center. Suppose the random variables ξ_i under the null hypothesis are independent and has the Poisson probability distribution with a parameter (expected value) $\mu_i, i = 1, \dots, m$:

$$H_0 : E\{\xi_i\} = \mu_i, \xi_i \sim Pois(\mu_i), i = 1, \dots, m.$$

Under Poisson model let us define a local statistic for each district i :

$$z_i = \frac{c_i - \mu_i}{\sqrt{\mu_i}},$$

where μ_i denote the expected number of cases respectively under the null hypothesis of not increased risk in district i .

Expected number of cases can be calculated as:

$$\mu_i = \frac{C}{N} n_i.$$

Under the null hypothesis statistic z_i^2 can be considered as a random variable with chi-square distribution with one degree of freedom and distribution function $F_{\chi^2_1}(x)$.

Consider defining a statistic:

$$M = \max_i z_i^2. \quad (1)$$

M statistic under the null hypothesis has distribution

$$F_M(x) = P\{M \leq x\} = P\{z_1^2 \leq x, z_2^2 \leq x, \dots, z_m^2 \leq x\} = \prod_{i=1}^m P\{z_i^2 \leq x\} = F_{\chi_1^2}^m(x).$$

Then the hypothesis test for global clustering is constructed:

$$\text{Choose hypothesis } \begin{cases} H_0, & \text{if } P > \alpha, \\ H_1, & \text{if } P \leq \alpha, \end{cases}$$

where P -value is calculated as $P = 1 - F_{\chi_1^2}^m(M)$, α is a significance level.

2. Cluster construction using the spatial scan statistic

Let us define a window as a set of districts and construct it in the following way. Impose a circle on every district with district's centroid being a center of the circle. The radius of the circle varies from 0 to a pre-defined maximum distance d , or up to the value at which the circle contains the maximum allowed number of districts K . If a window contains a district's centroid, then this district is fully included in the window. We are ending up with very large number of different but overlapping circular windows, each with different location. Every such window is a potential cluster.

The size of a circular window could be limited in order not to include more than 50% of the total population at risk. This limitation allows detecting clusters of both small and large sizes.

The most important aspect in the model construction for the scan statistic method realization is the choice of the distribution under the null hypothesis on the absence of clusterization. Usually Poisson or binomial probability distributions are used. To calculate the expected number of cases in a district the data on the population of the district or on the number of the risk group are used.

Let $Z_{ik}, k = 1, \dots, K$ denote a window which contains $(k - 1)$ nearest neighbors of i -th district. All windows to be scanned with circular spatial scan statistic are included in set [2]:

$$Z_1 = \{Z_{ik} \mid 1 \leq i \leq m, 1 \leq k \leq K\}. \quad (2)$$

The spatial scan statistic is constructed in accordance with the likelihood ratio test [2]:

$$S = \sup_{Z \in Z_1} \lambda(Z) = \sup_{Z \in Z_1} \left(\frac{c_Z}{\mu_Z} \right)^{c_Z} \left(\frac{C - c_Z}{C - \mu_Z} \right)^{C - c_Z} I \left(\frac{c_Z}{\mu_Z} > \frac{C - c_Z}{C - \mu_Z} \right), \quad (3)$$

where $c_Z = \sum_{i \in Z} c_i$ is the observed number of incidence cases in the window Z , $\mu_Z = \sum_{i \in Z} \mu_i$ is the expected number cases in the window Z , $I(\cdot)$ is the indicator function.

The window $Z^* \in Z_C$ that gives the maximal value to the statistic (3) is the cluster searched for with the highest probability value. The procedure of the statistical significance testing is organized with the use of the Monte-Carlo method.

Monte Carlo hypothesis testing is a 4 step procedure:

- 1) Calculate test statistic on real data.
- 2) Create a large number of random datasets generated under null hypothesis.
- 3) Calculate test statistic for every simulation.
- 4) Sort the resulting real and simulated data statistic values and mark the position of the statistic value which was calculated on the real dataset. If it appears at the top α portion of the whole set, then the null hypothesis is rejected with significance level α .

The algorithm of the spatial scan statistic can be modified to analyze spatio-temporal data. In this case, the time is introduced as the third measurement (coordinate), and the circular windows for calculation of the spatial scan statistic are replaced by cylinders. The base of these cylinders corresponds to some area, as in the spatial case, and the height means the spread of the potential cluster in time. The formula for the statistic (3) calculation remains the same, but the set of scanning windows is changed. The windows turn to cylinders, as the third (time) coordinate is added. The set Z_C in the formula (1) is replaced by the set Z_{ST} :

$$Z_{ST} = \{Z_{ik.[a,b]} \mid 1 \leq i \leq m, 1 \leq k \leq K; a, b = \overline{T_1, T_p}, a \leq b\},$$

where $Z_{ik.[a,b]}$ means the window of the cylinder form, including the district i and its $(k-1)$ nearest neighbors for each time interval T_p from the set $\{T_a, T_{a+1}, \dots, T_b\}$.

3. Robust version of the spatial scan statistic construction

If the probability model describes observations with outliers, we consider the so-called robust statistical methods of the spatial cluster analysis. Let \bar{c}_z be the sample mean of cases in window Z , and $|Z|$ be the number of cases in window Z . Analogously, let \bar{C} be the sample mean of all cases, and $|C|$ be the total number of cases. As $c_z = \bar{c}_z |Z|$ and $C = \bar{C} |C|$, expression (3) may be written in the form:

$$S = \sup_{Z \in Z_1} \lambda(Z) = \sup_{Z \in Z_1} \left(\frac{\bar{c}_z |Z|}{\mu_Z} \right)^{\bar{c}_z |Z|} \left(\frac{\bar{C} |C| - \bar{c}_z |Z|}{\bar{C} |C| - \mu_Z} \right)^{C - \bar{c}_z |Z|} I \left(\frac{\bar{c}_z |Z|}{\mu_Z} > \frac{\bar{C} |C| - \bar{c}_z |Z|}{\bar{C} |C| - \mu_Z} \right). \quad (4)$$

If data has at least one outlier, then statistic (3) often determine the cluster that include only this outlier. If the goal of our research is to find a cluster spread in space or time, then we need to determine the lower bound of number of observation in the cluster for reducing outlier influence.

As under outliers \bar{c}_z is a biased estimator for the location parameter, the due to a sufficient number of observation in a cluster, the robust estimator of the mean was used in (4) instead of \bar{c}_z . The robust estimators of the mean proposed by Hampel, Andrew's, Huber and Winsor's mean [4] were used. A spatial scan statistic sensitivity to outliers in the cluster was analyzed by using robust estimators of the mean [5].

4. Cluster construction using the flexible spatial scan statistic

An important parameter for the use of the spatial scan statistic method is the form of the scanning window. The use of round windows makes not possible to find clusters of the special forms, e.g., long areas along the rivers, parts of the polluted territories.

We will consider a method based on a statistic which allows detecting clusters with different flexible shapes [3]. These clusters will be limited with relatively small number of nearest neighbors of each district.

Let $Z_{ik}, k = 1, \dots, K$ denote a window which contains $(k-1)$ nearest neighbors of i -th district.

The flexible scan statistic imposes an irregularly shaped window Z on every district by connecting its neighbor districts. For every district i we create a set of windows with irregular shapes which contain k connected districts including district i . Let k vary from 1 to a pre-set maximum K . To avoid detecting clusters with unlikely peculiar shapes, we limit connected districts as a subset of set of district i and its $(K-1)$ nearest neighbors, where K is a pre-set maximum length of cluster. Alike to the circular spatial scan statistic a large number of different but overlapping windows with irregular shapes are created.

Let $Z_{ik(j)}, j = 1, \dots, j_{ik}$ denote the j -th window which is a set of k connected districts starting from district i , where j_{ik} is the number of j values meeting condition $Z_{ik(j)} \subseteq Z_{ik}, k = 1, \dots, K$. As the result all windows to be scanned are included in set

$$Z_2 = \{Z_{ik(j)} \mid 1 \leq i \leq m, 1 \leq k \leq K, 1 \leq j \leq j_{ik}\}.$$

For each district i the circular spatial scan statistic scans K concentric circles while the flexible scan statistic scans K concentric circles plus all sets of connected districts (including district i) with centers lying within the K -th largest concentric circle.

5. Clustering of thyroid carcinoma cases

A cluster detection study was made for thyroid carcinoma diagnostic data among children up to 18 years old from 1989 to 2005 in the Republic of Belarus. Population at risk and number of cases was available for every year and district of the country. 98.699 km (is equivalent to 1° of eastern longitude for Belarus) and 111.272 km (is equivalent to 1° of northern latitude for Belarus) constants were used for calculating distances in kilometers. The whole territory of the country is divided into 119 districts, the population and the incidence data for every district were used in the analysis. Administrative centroid coordinates will be used to determine the geographical position of every district.

Test based on statistic (1) was used for investigation of global clusterization. A global clusterization was detected at the level of significance $\alpha = 0.05$ in 1991-2003 years.

The thyroid carcinoma diagnostical dataset was analyzed using the method of the spatial scan statistic with maximum cluster size set to $K = 20$. The data for each year were analyzed separately, the dependence on data for other years was omitted. The 13 statistically significant clusters were detected among thyroid carcinoma cases. P -values for each cluster were estimated using 999 Monte Carlo simulations.

The results are given in table. Every cluster is represented by its centroid, the number of contained districts, the number of cases inside the cluster and the p -value.

The thyroid carcinoma clusters detected by the method of the spatial scan statistic

Year	Centroid	Districts	Cases	P -value
1991	Brahin	13	38	0.000
1992	Rečyca	12	24	0.000
1993	Stolin	5	24	0.000
1994	Brahin	15	46	0.000
1995	Lojeŭ	12	44	0.000
1996	Rečyca	15	40	0.000
1997	Lojeŭ	9	33	0.000
1998	Stolin	1	9	0.000
1999	Brahin	15	41	0.000
2000	Khojniki	14	35	0.000
2001	Lojeŭ	11	27	0.000
2002	Stolin	4	16	0.000
2003	Lojeŭ	13	13	0.040

Clusters detected using the spatio-temporal scan statistic are presented in fig. 1.

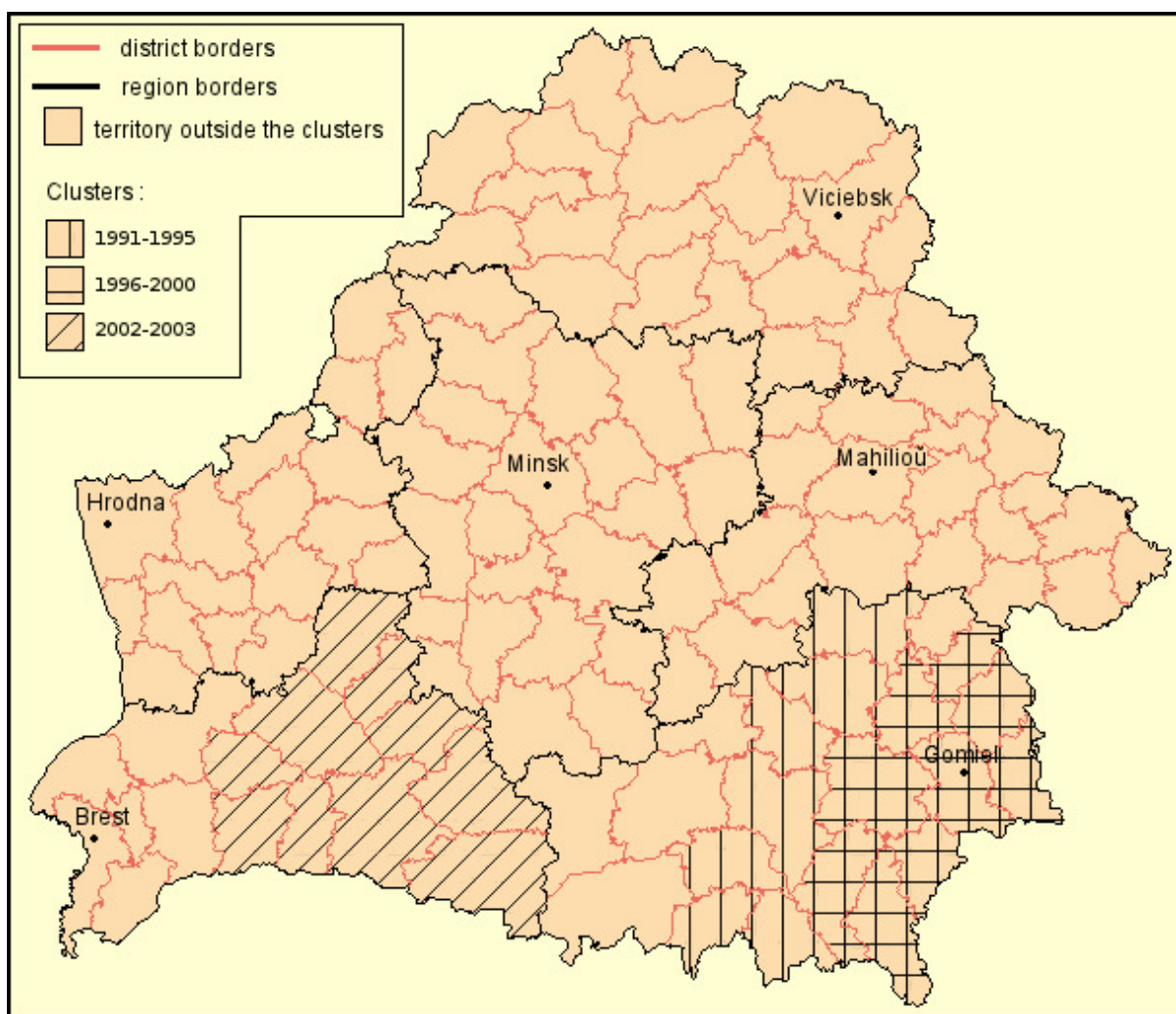


Fig. 1. Thyroid carcinoma clusters detected using the method of the spatio-temporal scan statistic

Conclusion

Spatio-temporal cluster analysis confirmed that there was a significant increase in the number of thyroid carcinoma cases among children aged from 0 to 18 throughout the territory of Gomel region in the 1990s and Brest region in the 2000s. These investigations were supported by ISTC (Project B-1910).

References

1. Rogerson, P. A set of associated statistical tests for detection of spatial clustering / P. Rogerson // *Ecological and Environmental Statistics*. – 2005. – Vol. 12. – P. 275–288.
2. Tango, T. A spatial scan statistic with a restricted likelihood ratio / T. Tango // *Japanese Journal of Biometrics*. – 2008. – Vol. 29, №. 2. – P. 75–95.
3. Tango, T. A flexibly shaped scan statistic for detecting clusters / T. Tango, K. Takahashi // *International journal of Health Geographics*. – 2005. – Vol. 4. – P. 115– 125.
4. Huber, P.J. *Robust Statistics* / P.J. Huber. – N.Y. : Wiley, 1981.
5. Abramovich, M.S. Robust spatio-temporal cluster analysis of disease / M.S. Abramovich, M.M. Mitskevich // *Proceedings of the 10th Intern. Conference “Computer Data Analysis and Modeling”*. – Vol. 2. – Minsk : Publishing center of BSU. – 2013. – P. 95–98.

MODELING OF FLOW DISTRIBUTION IN ELECTRIC NETWORK TAKING INTO ACCOUNT THE UNCERTAIN FACTORS

O. Alexandrov, D. Svirsky, N. Radoman
Belarusian State Technology University, Minsk
e-mail: d_svir@mail.ru

Solving of the problem of a flow distribution in the power system at absence or uncertainty of missing information is considered in the report. Algorithms of calculation of uncertain parameters of the electrical power system established mode are offered.

Introduction

Calculation of the established mode of an electric network of electrical power system (EPS) can be made in the presence of full information on parameters of the EPS elements (resistance, conductivity, coefficients of transformation of transformer branches, etc.) and independent parameters of a mode (the nodal active and reactive capacities arriving from knot in a network). Other (dependent) parameters (power streams in branches, voltage in knots) are defined by a settlement way [1, 2]. Information on current state of EPS is determined by data on the mode parameters, arriving in the form of telesignals by TV channels from the corresponding sensors. Currents, capacities and tension modules are measured in knots usually. Therefore instead of the full vector of the nodal capacities used as entrance information for calculation of the established mode, there are partially certain vectors of streams of power in branches, nodal capacities and tension in knots. If we consider that instruments of devices are read at the same time, the measured parameters of a mode have to be the task to count unknown parameters connected among themselves by the relevant laws of Ohm and Kirchhoff. Let's set the task to count unknown parameters of a mode using all set of made telemeasurements. It is possible to offer for this purpose two approaches.

1. The task to count unknown parameters

Let's write down two equations of a condition in which all measured parameters are connected among themselves by turned form:

$$\begin{aligned}U_{\Delta}' &= Z'J + DE, \\I' &= CJ + YE,\end{aligned}$$

where U_{Δ}' – vector-column of voltage in knots of rather basic knot; Z' – matrix of nodal resistance of the scheme; J – vector-column of nodal currents; D – matrix of coefficients of distribution of voltage; E – EMF column-vector in branches; I' – column-vector of currents in branches; C – matrix of coefficients of distribution of currents; Y – matrix of own and mutual admittances in branches.

Let's consider in vectors U_{Δ}' and I' the additional tension and the currents caused by existence of EMF in branches:

$$U_{\Delta} = U_{\Delta}' - DE; \quad I = I' - YE,$$

then

$$U_{\Delta} = Z'J; \tag{1}$$

$$Z = CJ. \tag{2}$$

Let's pass from currents to capacities, having increased both parts of the equations (1) and (2) by the average voltage of a network U .

Then receive

$$U_{\Delta} = \frac{1}{U} Z's = Zs, \tag{3}$$

where

$$Z = \frac{1}{U} Z' \quad S = Cs. \quad (4)$$

As nodal capacities are measured not in all knots of a network, the vector S with isn't defined, and direct calculations on formulas (3) and (4) can't be carried out.

Let's break vector s into blocks: $s = [s^+ | s^-]^*$ where S^+ – the block consisting of known components of the vector S ; S^- – the block consisting of unknown components of the vector S .

Let's break vectors U_Δ and S similarly:

$$U_\Delta = [U_\Delta^+ | \mathbf{M}_\Delta^-]; \quad S = [S^+ | \mathbf{M}^-].$$

Let's break into blocks as well matrixes Z and C :

$$Z = \left[\begin{array}{c|c} Z^{++} & Z^{+-} \\ \hline Z^{-+} & Z^{--} \end{array} \right]; \quad C = \left[\begin{array}{c|c} C^{++} & C^{+-} \\ \hline C^{-+} & C^{--} \end{array} \right].$$

Then the equations (3) and (4) will correspond in a look:

$$\begin{cases} U_\Delta^+ = Z^{++} s^+ + Z^{+-} s^-; \\ U_\Delta^- = Z^{-+} s^+ + Z^{--} s^-; \end{cases} \quad (5)$$

$$\begin{cases} S^+ = C^{++} s^+ + C^{+-} s^-; \\ S^- = C^{-+} s^+ + C^{--} s^-. \end{cases} \quad (6)$$

Let's transform the first equations of systems as follows:

$$U_\Delta^+ - Z^{++} s^+ = Z^{+-} s^-; \quad (7)$$

$$S^+ - C^{++} s^+ = C^{+-} s^-. \quad (8)$$

Let's unite the equations (7) and (8) in system and we will designate:

$$A = \left[\begin{array}{c} U_\Delta^+ - Z^{++} s^+ \\ S^+ - C^{++} s^+ \end{array} \right]; \quad B = \left[\begin{array}{c} Z^{+-} \\ C^{+-} \end{array} \right].$$

Taking into account it we will receive:

$$A = Bs^-. \quad (9)$$

Can receive the decision of this system one of methods of calculus mathematics, for example, by means of procedure of the conversion of the matrix B (if it is quadratic): $s^- = B^{-1} A$.

System (9) not degenerate if the number of unknown is equal to number of the equations of system. From here the necessary condition of existence of the decision of system of the linear equations (9) follows. The sum of number of measurements of tension in knots and streams in branches isn't less missing number of measurements of nodal capacities. It should be noted that the conversion of the matrix B difficult as it is densely filled with complex numbers.

Generally the matrix B isn't quadratic and the vector s^- will be the normal pseudo-decision of expression (9)

$$s^- = B^- A, \quad (10)$$

where B^- – Moore-Penrose's pseudo-return matrix [2] in relation to the matrix B. In cases when (9) has the only decision, $B^- = B^-$. For definition of the pseudo-return matrix there are special algorithms which have, however, a number of computing difficulties [3].

The vector s^- , defined from the equation (10), supplements the vector s^+ to the full vector s , then the usual problem of calculation of the flow distribution in difficult closed system of EPS is solved.

2. Design is based on equations of a condition

The second approach is based on record of the equations of a condition of an electric network in a direct form. Without EMF in branches of the scheme they can be written down as follows:

$$MI = J, \quad (11)$$

$$YU_{\Delta} = J, \quad (12)$$

where M – the first connection matrix of the scheme; Y – the matrix of nodal conductivity of the scheme. The vectors J, I, U_{Δ} components are partially certain as well as at the first approach that also gives the chance to carry out direct calculations on expressions (11), (12).

Let's break matrixes entering into formulas and vectors into blocks:

$$M = \left[\begin{array}{c|c} M^{++} & M^{+-} \\ \hline M^{-+} & M^{--} \end{array} \right]; \quad Y = \left[\begin{array}{c|c} Y^{++} & Y^{+-} \\ \hline Y^{-+} & Y^{--} \end{array} \right];$$

$$U_{\Delta} = \left[\begin{array}{c} U_{\Delta}^+ \\ \hline U_{\Delta}^- \end{array} \right]; \quad I = \left[\begin{array}{c} I^+ \\ \hline I^- \end{array} \right]; \quad J = \left[\begin{array}{c} J^+ \\ \hline J^- \end{array} \right],$$

where the index (+) belongs to measuring components and (-) - to not measured. Then the equations (11) and (12) will correspond in a look:

$$\begin{cases} M^{++}I^+ + M^{+-}I^- = J^+; \\ M^{-+}I^+ + M^{--}I^- = J^-; \end{cases} \quad (13)$$

$$\begin{cases} Y^{++}U_{\Delta}^+ + Y^{+-}U_{\Delta}^- = J^+; \\ Y^{-+}U_{\Delta}^+ + Y^{--}U_{\Delta}^- = J^-. \end{cases} \quad (14)$$

As components with an index (+) are known that it is possible to divide known and unknown variables from the equations (13) and (14):

$$M^{+-}I^- = J^+ - M^{++}I^+; \quad (15)$$

$$Y^{+-}U_{\Delta}^- = J^+ - Y^{++}U_{\Delta}^+. \quad (16)$$

If the number of unknown doesn't exceed number of the equations of system matrix equations are not degenerate. From here the necessary condition for definition of components of vectors I and U_{Δ}^- follows: the number of missing measurements of tension and currents in branches shouldn't surpass number of nodal measurements of currents.

In respect of a problem of arrangement of telemeasurements in EPS it means that the preference should be given to block measurements when tension is measured in knot, and an injection of nodal power, and currents in departing lines also.

Let's consider in more detail allocation procedure of extraction from the matrix M the matrix M^- which is that the lines corresponding to knots are deleted from a matrix of M, not having measurement of injections, and the columns corresponding to branches, having measurements of currents. Such procedure can be modeled by means of a matrix of transformations; which turns out from the single matrix which order is equal to number of branches of the scheme by deletion of the corresponding lines, and movement

of the next lines up. Similarly occurs if it is necessary to remove a matrix M column. Then the column is deleted, and other columns move to the left.

In a matrix look during removal of the line p we will receive:

$$M^- = \Phi_p M,$$

where M^- – the M matrix after deletion of the line p ; Φ_p – corresponding matrix of transformations; p – number of a deleted line. Then the conversion of the matrix M can be replaced with the conversion of the product $(\Phi_p M)$, i.e.

$$(M')^{-1} = (\Phi_p M)^{-1}$$

or

$$(M')^{-1} = M^{-1} \Phi_p^{-1}.$$

Obviously, $\Phi_p^{-1} = \Phi_p^*$, as

$$\Phi_p^{-1} \Phi_p = \Phi_p^* \Phi_p = 1,$$

where 1 – a single matrix and therefore last expression will correspond in a look:

$$(\Phi_p M)^{-1} = M^{-1} \Phi_p^*$$

or, identically:

$$(\Phi_p M)^{-1} = C \Phi_p^*.$$

From this the conclusion follows: the conversion of the matrix M with remote line is equivalent to removal of the corresponding column of the matrix of coefficients of distribution. The similar conclusion can be drawn concerning a matrix M column, i.e. it is equivalent to removal of a line of the matrix C . This result can be extended to any number of excluded lines and matrix M columns, and also to the matrix Y . Having united procedure of removal of lines with procedure of removal of columns, we will receive:

$$M^{+-} = \Phi^+ M T^-,$$

where Φ^+ – matrix of transformations in which remote there were corresponding lines; T^- -matrix of transformations in which the corresponding columns are deleted. Then

$$(M^{+-})^{-1} = (\Phi^+ M T^-)^{-1} = (T^-)^* C (\Phi^+)^*$$

or

$$(M^{+-})^{-1} = C^{-+}. \quad (17)$$

Similarly

$$(Y^{+-})^{-1} = Z^{-+}. \quad (18)$$

Substituting (17) and (18) in (15) and (16) accordingly, we will receive:

$$I^- = C^{-+} J_1^+,$$

$$U_{\Delta}^- = Z^{-+} J_2^+.$$

The received expressions remove the uncertainty of the parameters in formulations 11 and 12.

Conclusion

Thus it is possible to count unknown parameters of a mode without resorting to flow distribution calculation in a network by means of last formulas at observance necessary condition of an undegeneracy of the equations (1) (2). As the algorithms of joint formation and operative correction of matrixes Z and C at commutation in a network are known, this approach allows counting unknown parameters without resorting to procedures of the matrixes conversion owing to that the efficiency of the offered algorithm is extremely high.

References

1. Hazra, J. Congestion management using multiobjective particle swarm optimization / J. Hazra, A.K. Sinha // IEEE Transactions on Power Systems. – 2007 – Vol. 22 (4). – P. 1726–1734.
2. Vijayakumar, K. Multiobjective Optimization Methods for Congestion Management in Deregulated Power Systems / K. Vijayakumar // Journal of Electrical and Computer Engineering. – 2012. – Vol. 2012. – P. 1–8.
3. Алберт, А. Регрессия, псевдоинверсия и рекуррентное оценивание / А. Алберт. – М. : Наука, 1977. – 223 с.
4. Гантмахер, Ф.Р. Теория матриц / Ф.Р. Гантмахер. – М. : Наука, 1988. – 548 с.

STATE SPACE REPRESENTATION OF MULTI-INPUT MULTI-OUTPUT CIRCUITS

Batmunkh Amar
Computer science department
School of Computer Science and Management
Mongolian University of Science and Technology
e-mail: batmunkh@csms.edu.mn

This article considers conversion of transfer function of a system to state space representation in concept of MIMO system. As relatively real system there was chosen universal filter circuit and it's simulation in MATLAB.

Introduction

Although the idea of state space representation of a system behavior was known since 1960 it's application becomes widely last few decades [1]. The adventures of the state space representation of dynamic system or process becomes obviously in analyzing of multi-input and multi-output (MIMO) systems and it shows that state space method becomes one of the widely used classical methods for modelling and controlling any process using state variables as internal arguments. Also any mathematically modellable process or system can be mapped from continuous time domain into discrete time domain through Z transform [2, 3]. This paper attempts to show an approach to combine the universal filter and it's state space representation.

1. Methods used

In this section will be discussed used methods and objects that are under consideration.

A. State space representation

State space form has own advantages relatively to other methods [4]. Any transfer function can be converted into state space representation in the form of (1) and this concept is also can be propagated for MIMO systems

$$\begin{aligned} \dot{x} &= Ax + Bu; \\ y &= Cx + Du, \end{aligned} \quad (1)$$

where A, B, C, D are matrices having sizes $n \times n$, $n \times p$, $m \times n$ and $m \times p$ and n – number of states, m – number of system outputs, p – number of system inputs. (1) shows the system is in continuous time domain. When transfer functions converted to state space form equation (1) is converted to (2) in Laplace domain [4]

$$\begin{aligned} sX(s) &= AX(s) + BU(s); \\ Y(s) &= CX(s) + DU(s). \end{aligned} \quad (2)$$

There are many articles related to building of discrete time representation of state space method and it is used for feedback discrete control systems based on feedback law and state feedback observer. Common view of discrete state space representation is shown in (3)

$$\begin{aligned} \hat{x}_{k+1} &= A_d \hat{x}_k + B_d u_k; \\ y_k &= C_d \hat{x}_k + D_d u_k, \end{aligned} \quad (3)$$

where A_d, B_d, C_d and D_d are Z-transformed matrices and \hat{x}_k is discrete state and u_k is discrete value of input excitation.

B. Multi-input multi-output system

In some cases discussed before when system has multi inputs and multi outputs the best method for analyzing is state space and the methods for analyzing SISO not always applicable for MIMO [4]. As described in [5] conversion of transfer function of MIMO system can be stated

in 8 steps, that could be bundled into following common concepts as choosing corresponding column and finding common factors, using intermediate variable to note additional functions and getting out number of independent expressions in state space form. As an example, let us consider system or plant of active filter on operational amplifiers so called universal filter.

C. Universal filter

Schematic of universal filter can be found in many texts describing analogue filters. There was chosen schematic shown in fig.1 for simplicity [6]. It has one input and three outputs (HPF, BPF-resonance, LPF) so it can be seen as MIMO system. Applying the node voltage method to the circuit in frequency domain one can get the following transfer functions.

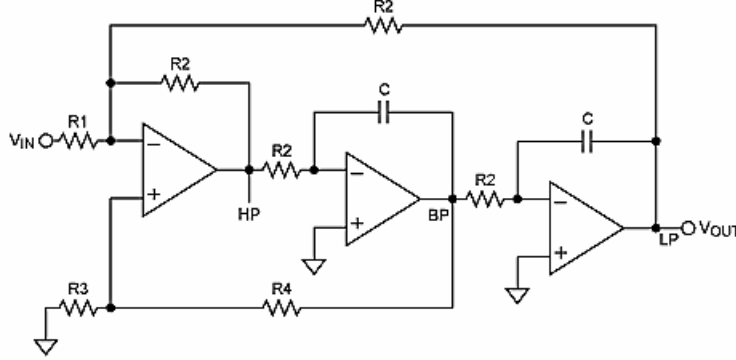


Fig. 1. Active universal filter

From fig. 1 transfer function for each filter is

$$K_{hp}(s) = -\frac{R_2}{R_1} \times \frac{s^2 R_2^2 C^2}{1 + s(R_2 + 2R_1)R_2 R_3 C / R_1(R_3 + R_4) + s^2 R_2^2 C^2}; \quad (4)$$

$$K_{lp}(s) = -\frac{R_2}{R_1} \times \frac{1}{1 + s(R_2 + 2R_1)R_2 R_3 C / R_1(R_3 + R_4) + s^2 R_2^2 C^2}; \quad (5)$$

$$K_{bp}(s) = \frac{R_2}{R_1} \times \frac{s R_2 C}{1 + s(R_2 + 2R_1)R_2 R_3 C / R_1(R_3 + R_4) + s^2 R_2^2 C^2}. \quad (6)$$

2. State space representation of the object

Using method proposed in [5] the transfer functions can be converted into following form

$$\begin{bmatrix} U_{hp}(s) \\ U_{lp}(s) \\ U_{bp}(s) \end{bmatrix} = \frac{1}{1 + s(R_2 + 2R_1)R_2 R_3 C / R_1(R_3 + R_4) + s^2 R_2^2 C^2} \times \begin{bmatrix} -s^2 R_2^2 C^2 \frac{R_2}{R_1} \\ -\frac{R_2}{R_1} \\ s R_2 C \frac{R_2}{R_1} \end{bmatrix} U_{in}(s). \quad (7)$$

Using additional intermediate variable W to denote the common multiple and including input excitation u_{in} we get

$$W(s) = \frac{1}{1 + s(R_2 + 2R_1)R_2 R_3 C / R_1(R_3 + R_4) + s^2 R_2^2 C^2} U_{in}(s) \quad (8)$$

or

$$W(s)(1 + s(R_2 + 2R_1)R_2 R_3 C / R_1(R_3 + R_4) + s^2 R_2^2 C^2) = U_{in}(s). \quad (9)$$

In time domain last expression (9) becomes (10):

$$w + \dot{w}(R_2 + 2R_1)R_2R_3C/R_1(R_3 + R_4) + \ddot{w}R_2^2C^2 = u_{in}(t) \quad (10)$$

or

$$\ddot{w} = -\frac{1}{R_2^2C^2}w - \frac{(R_2 + 2R_1)R_2R_3C}{R_2^2C^2R_1(R_3 + R_4)}\dot{w} + \frac{1}{R_2^2C^2}u_{in}(t) \quad (11)$$

and using (7), (11) and after some manipulations the outputs are

$$u_{hp}(t) = -\frac{R_2}{R_1}R_2^2C^2 \left(-\frac{1}{R_2^2C^2}w - \frac{(R_2 + 2R_1)R_2R_3C}{R_2^2C^2R_1(R_3 + R_4)}\dot{w} + \frac{1}{R_2^2C^2}u_{in}(t) \right); \quad (12)$$

$$u_{lp}(t) = -\frac{R_2}{R_1}w; \quad (13)$$

$$u_{bpf}(t) = \frac{R_2}{R_1}R_2C\dot{w}. \quad (14)$$

Using regular notation of state space representation of variables $x_1 = w$, $x_2 = \dot{x}_1 = \dot{w}$, $\dot{x}_2 = \ddot{x}_1 = \ddot{w}$, above expressions have been modified into following matrix form

$$\left\{ \begin{array}{l} \dot{x}_1 = x_2; \\ \dot{x}_2 = -\frac{1}{R_2^2C^2}x_1 - \frac{(R_2 + 2R_1)R_2R_3C}{R_2^2C^2R_1(R_3 + R_4)}x_2 + \frac{1}{R_2^2C^2}u_{in}; \\ y_{hp} = \frac{R_2}{R_1}x_1 + \frac{(R_2 + 2R_1)R_2^2R_3C}{R_1^2(R_3 + R_4)}x_2 - \frac{R_2}{R_1}u_{in}; \\ y_{lp} = -\frac{R_2}{R_1}x_1; \\ y_{bpf} = \frac{R_2}{R_1}R_2Cx_2. \end{array} \right. \quad (15)$$

(15) shows that this system has two state variables and moreover the matrix D does't equals to zero. From (15) the system state matrices of our system as follows

$$A = \begin{bmatrix} 0 & 1 \\ -1/R_2^2C^2 & -\frac{(R_2 + 2R_1)R_3}{R_2C R_1(R_3 + R_4)} \end{bmatrix}; \quad B = \begin{bmatrix} 0 \\ 1/R_2^2C^2 \end{bmatrix};$$

$$C = \begin{bmatrix} R_2/R_1 & \frac{(R_2 + 2R_1)R_2^2R_3C}{R_1^2(R_3 + R_4)} \\ -\frac{R_2}{R_1} & 0 \\ 0 & \frac{R_2^2C}{R_1} \end{bmatrix}; \quad D = \begin{bmatrix} -\frac{R_2}{R_1} \\ 0 \\ 0 \end{bmatrix}.$$

Frequently systems do not have the term D and in most examples related to this problem $D = 0$. But in this case it does not equals to zero. It means that high frequency output of the circuit directly depends on input as shown above.

From denominator of (4), (5) or (6) the resonance frequency (for BPF output) or 3dB frequency (for HPF and LPF outputs) is defined as $f_{3dB} = (2\pi R_2C)^{-1}$, filter gain $K_p = R_2/R_1$, and quality factor $Q = R_1(R_3 + R_4)/[R_3(R_2 + 2R_1)]$.

To simulate above obtained system of transfer functions of MIMO system in state space form let be $R_1 = 1 \text{ k}\Omega$, $R_2 = 4 \text{ k}\Omega$, $C = 1 \mu\text{F}$, $R_3 = 2 \text{ k}\Omega$, $R_4 = 10 \text{ k}\Omega$. Although to design any filter first of all should be set value of half power frequency, quality factor and gain of the filter. But in this case for the above arbitrary chosen values substituting them into corresponding expressions above, consequently, we get resonance frequency or $f_{3dB} = 39.8 \text{ Hz}$; $K_o = 4$; $Q = 1$.

3. Chirp signal

Now, let consider the input excitation, as an example, in the form of linear chirp signal due to it's dynamic behavior in time domain which's time expression is (16).

$$chirp(t) = A \sin[\varphi_0 + 2\pi(f_0 + kt)t], \quad (16)$$

where A is amplitude, φ_0 is initial phase, f_0 is initial frequency, k is rate coefficient (it has unit $1/\text{sec}^2$) and t is time. From (16) we get $f = f_0 + kt$ or - the frequency dynamics of the signal and $k = f - f_0/t$.

To show behavior of the filter, were chosen frequency range from $f_0 = 0 \text{ Hz}$ to $f_{final} = 90 \text{ Hz}$ with zero phase and unit amplitude and timing interval of 10 seconds. Taking all these into consideration below shown calculation steps of the process.

Because $f_0 = 0$ and simulation final time $t_{final} = 10 \text{ sec}$ the rate coefficient $k = 90/10 = 9$. Also, the time instant, at which amplitude of the output signal on BPF output of the filter reaches it's maximum value (resonance), is calculated as follows,

$$t_{resonance} = \frac{f_{3dB} \times t_{final}}{f_{final}} = \frac{39.8 \text{ Hz} \times 10 \text{ sec}}{90 \text{ Hz}} = 4.42 \text{ sec}.$$

4. Simulation in MATLAB

Three experiments were conducted using either *Simulink*, *SimScape* and *lsim(sys, u, t)* function.

A. Command lines below create state matrices and simulate state space model of our system "filter".

```
t=0:0.001:10;
u=chirp(t,0,tf,90,'linear');
a=[0 1; -1/R2^2/C^2 -R3*(R2+2*R1)/R2/C/R1/(R3+R4)];
b=[0; 1/R2^2/C^2];
c=[R2/R1 (R2+2*R1)*R2^2*R3*C/R1^2/(R3+R4); -R2/R1 0; 0 R2^2*C/R1];
d=[-R2/R1; 0; 0];
filter=ss(a,b,c,d);
y=lsim(filter,u,t);
plot(t,y);
```

B. Fig. 2 and 3 show the SimScape and Simulink simulation diagrams for the circuit of fig.1.

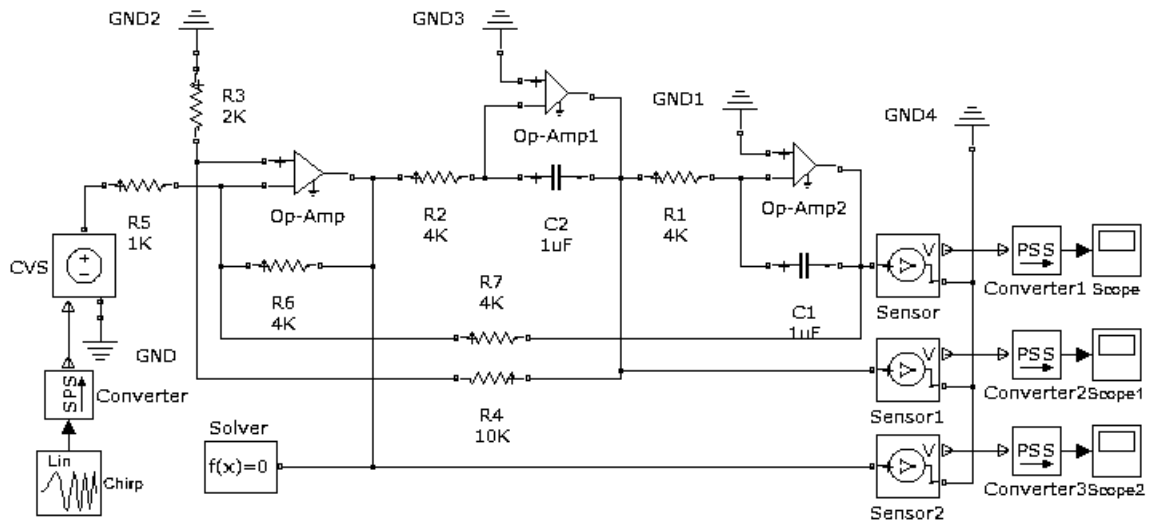


Fig. 2. Experiment of universal filter in SimScape

C. In fig. 4 shown three output timing waveforms of the filter on single graph. Results from the experiments demonstrated the same timing diagrams and because chosen time interval was 10 seconds and frequency range was 0-90Hz it is seen that the resonance frequency 39.8Hz corresponds at 4.42 sec as calculated before.

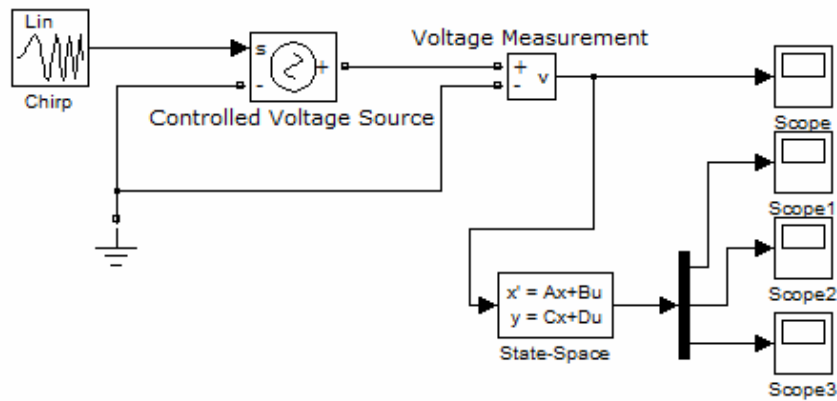


Fig. 3. Experiment of universal filter in Simulink

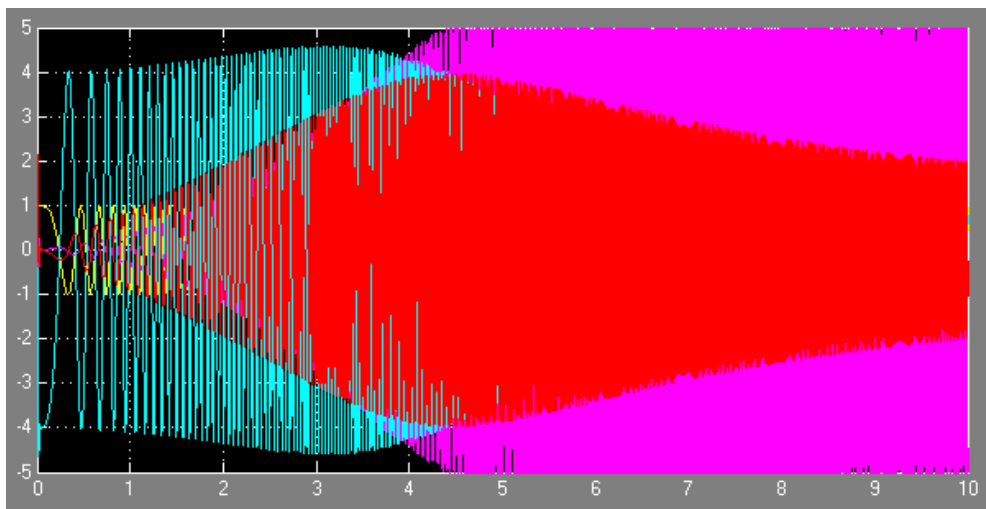


Fig. 4. Filter output three timing signals. Cyan - LPF, red-BPF and pink-HPF. It is seen that at 4.42 sec output of BPF is maximum (resonance) and because $Q=1$ the output of LPF slightly goes up in amplitude near f_{rdb}

Conclusion

This paper considers brief concepts of MIMO system in state space representation and considers as an example of MIMO system active filter based on three operational amplifiers which have one input and three outputs. Shown that system transfer functions are converted to state space representation using recommended method. As input excitation for the circuit was chosen chirp signal due to its dynamic behavior. Also was demonstrated in MATLAB filtering process in time and were created output timing waveforms using LSIM, SIMULINK and SIMSCAPE.

References

1. Neculai, A. Modern control theory. Historical perspective / A. Neculai. – Research Institute for informatics, 2005. – 5 p.
2. Damen, Ad. Modern control theory / Ad. Damen. – October 2002. – P. 149–157.
3. Smith, R. Designing digital control systems / R. Smith // ECE 147b. Discrete-time equivalents. – P. 1–7. – State feedback. – P. 1–8.
4. Di Ruscio, D. System theory state space analysis and control theory. Lecture notes / D. Di Ruscio. – 2009. – 35 p.
5. Pota, R. Hemanshu MIMO systems – Transfer function to State space / Hemanshu R. Pota // IEEE Transaction on education. – 1996. – P. 97–99.
6. A Beginner's Guide to Filter Topologies // Maximum integrated [Electronic resource]. – Mode of access : <http://www.maximintegrated.com/app-notes/index.mvp/id/1762>. – Date of access : 5.05.2014.

COMPUTER MODEL FOR DISTINGUISHING ASIAN PEOPLE BY FACE

B. Bat-Erdene¹, Ts. Ganbat²

¹The University of the Humanities, Ulaanbaatar, Mongolia;

²The Computer Science and Management School, the Mongolian

University of Science and Technology, Ulaanbaatar

e-mail: basubaer@gmail.com

In this paper, we aim to create new effective computer model for distinguishing Asian people by their face using frontal face part color, size and distances based on comparative research about image processing and face recognition. It is difficult to determine face part size and distance due to image quality, lighting condition, rotation angle and facial emotion. Hence, first we need to detect face on the image then convert image into real input. After that we can determine image candidate's gender, face shape, key points and face parts. Finally, we will return the result, based on comparison of sizes and distances with the sample's measurement table database. While we were measuring samples, there were big differences between images by their gender and face shapes. Input images must be the frontal face image that has smooth lighting and does not have any rotation angel.

Introduction

Face recognition provides a great opportunity to creating useful application to identify humans for the security and immigration offices in some countries. Nowadays there are some effective algorithms for detecting face candidate region. Even though some only works on Caucasian, others detecting face region. About the detection of face candidate nationality, especially research on distinguishing Asian people by their face have not been researched yet. To make a decision to solve this problem, first we have collected a thousand frontal image from every country (Mongolian, Japanese, Chinese, and Korean) to create our sampling. Then we have measured all of them to establish our measurement table database. The measurement provides great chances to processing images with database. For further development, we are working on creating full database that grouped through the common facial features.

Model for distinguishing by face (fig. 1)

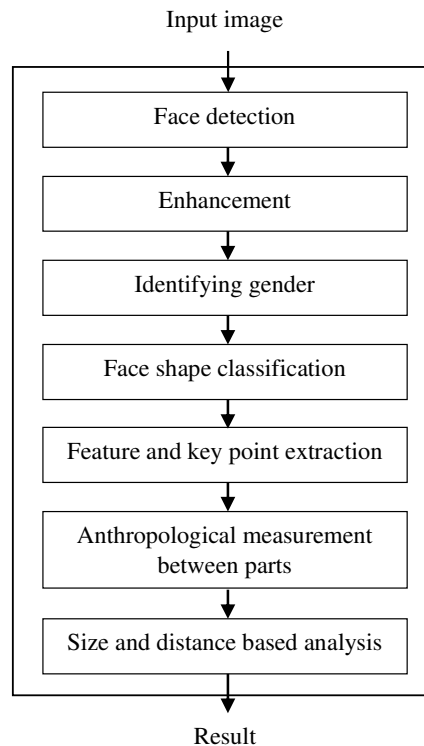


Fig. 1. Flowchart of the system model

Face detection

Face detection is a process to divide image into the two pieces: one contains face and the other non-face image. Final detection result has functional relation to the process time and overall result. A variety of methods has been proposed, and some have been brought into exploitation in real system.

In general, those methods can be divided into two major sections: the first section is based on all kind of face parts, which means the final result comes from the integration of several detection results. The second considers a face as a single detection input and facial features are extracted from the entire face region. We are going to use methods in both sections to achieve higher performance and better accuracy.

Enhancement

The better performance and accuracy are related to input lighting, rotation angle, facial expression and image quality. Hence, in this stage image size, quality lighting and rotation angle are normalized until real input.

Identifying gender

While we were taking measure from the sample images, there was interval exception from the image candidate's gender. In real biological systems, especially in an anthropological and morphological analysis calculator measurement must have been done with the gender classification.

Face shape classification

Human faces are categorized into seven geometric shapes. Using fig. 2's base shapes (from right oval, square, round, triangle, heart, oblong and diamond) we will identify the shapes from the facial image.



Fig. 2. Face shape categories

Feature and key point extraction

In our model, there are two types of facial features: base and extra features. The base facial features are the shapes and location of eyes, nose, mouth and lips. The extra features we have to detect are forehead, mandible, eyebrows, eyelid, chin and ear. We have measured all of these features by the color, size, shape and distance from the neighbors (fig. 3).

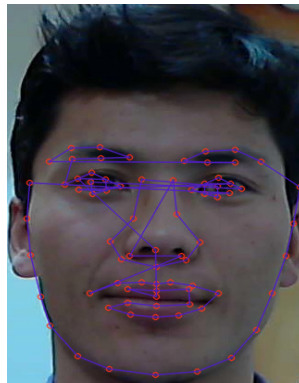


Fig. 3. Face detection key points

Anthropological measurement between parts

After feature and key point extraction, we calculated the distance from the neighbors using the landmark and edge detection. Then we will have full measurement collection from the image.

Size and distance based analysis

Above-mentioned measurements are created on frontal face images by the following values.

Size and distance table

№	Measurement units	
	Size	Distance
1	Head width, height	Eye distance
2	Forehead width, height	Eyebrow distance
3	Face width	Ear distance
4	Nose width, height	Chin to lip distance

5	Ear height	Lip to nose distance
6	Lip width	Nose to eye distance
7	Eye width, height	
8	Eyebrow width	
9	Chin width	

Conclusion

In this paper, we have presented a novel computer model for distinguishing Asian people by their face. First, a number of faces should be known before the process. Hence, we have collected and created our facial image database by the four countries.

To distinguish Asian people we have done two kind of feature extraction. It gives great opportunity to reduce similarities. The proposed model is effective for frontal face images with the normal lighting condition and with the normal rotation angle. Using this model, we can distinguish Asian people by single frontal image.

References

1. Lee, Ch.H. Fast Detection of Face Candidate for Passport Photographs / Ch.H. Lee // MITA. – 2012. – 108 p.
2. Bastanfard, A. Iranian Face Database with age, pose and expression / A. Bastanfard, M.A. Nik, M.M. Dehshibi // Machine Vision, 2007. International Conference ICMV 2007. – Dec. 2007. – P. 50–55.
3. Wang, J. A new face detection method based on shape information / J. Wang, T. Tan // Pattern Recogn. – Lett. 21. – June 2000. – P. 6–7.
4. Baker, J.R. Race / J.R. Baker. – N.Y. : London : Oxford University Press, 1974.
5. Vallois, H.V. Anthropometric Techniques / H.V. Vallois // Current Anthropology. – Vol. 6 (2). – 1965. – P. 127–143.
6. Shang, L. Temporal Exemplar-Based Bayesian Networks for Facial Expression Recognition / L. Shang, K.-P. Chan // Machine Learning and Applications, 2008. Seventh International Conference. – San Diego, California, USA, 2008. – P. 16–22.

INFORMATION EXTRACTION BASED ON SEMANTIC PATTERNS

A. Bobkov¹, S. Gafurov², V. Krasnoproshin¹, V. Romanchik², H. Vissia²

¹Belarusian State University, Minsk;

²Byelex Multimedia Products BV, Oud Gastel, The Netherlands

e-mail: anatoly.bobkov@gmail.com

The paper deals with information extraction from texts in a natural language. Special attention is paid to word collocations on the level of meaning and word sense disambiguation based on semantic patterns.

Introduction

Information extraction is one of the key elements in information technologies. The major part of the most interesting and important information can be found in a great variety of texts and documents in a natural language. Information extraction is gaining much popularity within natural language processing [1, 2]. The field of information extraction is well suited to various types of business and government intelligence applications. Diverse information is of great importance for decision making on products, services, persons, events, organizations.

Creation of systems that can effectively extract meaningful information requires overcoming a number of new challenges: identification of documents, knowledge domains, specific opinions, events, activities, as well as representation of the obtained results.

The purpose of this paper is to introduce an approach for solving the problem of effective extraction of meaningful information. Semantic patterns approach is proposed as a solution to the problem.

1. Problem statement and solution

Numerous models and algorithms are proposed for information extraction [3]. But the problem of effective information extraction from texts in a natural language still remains unsolved. Processing of texts in a natural language necessitates the solution of the problem of extracting meaningful information. Our approach is based on semantic patterns as main constituents in effective information extraction. The approach is mainly knowledge-driven.

In information extraction and text mining, word collocations show a great potential to be useful in many applications (machine translation, natural language processing, lexicography, etc.).

"Collocations" are usually described as "sequences of lexical items which habitually co-occur, but which are nonetheless fully transparent in the sense that each lexical constituent is also a semantic constituent" [4]. Such expressions as "fine weather", "torrential rain", "high winds", "cosmetic surgery" are examples of collocations.

The traditional method of performing automatic collocation extraction is to find a formula based on the statistical quantities of words to calculate a score associated to each word pair. Proposed formulas are mainly: "mutual information", "t-test", "z test", "chi-squared test" and "likelihood ratio" [5].

Word collocations from the point of semantic constituents have not yet been widely studied and used for extracting meaningful information, especially when processing texts in a natural language.

The proposed semantic patterns approach is based on word collocations on the semantic level and contextual relations. In general, a semantic pattern includes: 1) *participants* (a person, company, natural/manufactured object, as well as a more abstract entity, such as a plan, policy, etc.) involved in the action or being evaluated; 2) *actions* - a set of verb semantic groups and verbal nouns; 3) *rules for semantic patterns actualization*.

Examples of information extraction based on semantic patterns.

Example 1, Opinion Mining

The new generation Audi A8 employs active noise reduction in an attempt to offer a next-level quiet ride to its occupants.

Extracted instances:

Product = **Audi A8**

Product type = **Car**

Opinion words = **active noise reduction, next-level quiet ride**

Example 2, Economic Activities Detection

General manager Tim Deakin said Orkney Cheddar was produced with locally-sourced milk, following a traditional recipe and process.

Extracted instances:

Economic activity = ***Manufacture of Dairy Products***

Example 3, Event Extraction

A New York woman faced charges for faking cancer to solicit money from unsuspecting donors and a relative.

Extracted instances:

Event = ***Fraud & Forgery***

An ontology-based approach is used for semantic patterns actualization [6].

Ontologies have become common on the World-Wide Web [7]. Ontologies on the Web range from large taxonomies categorizing Web sites (such as on Yahoo!) to categorizations of products for sale and their features (such as on Amazon.com). For any given knowledge domain, the ontology represents the concepts which are held in common by the participants in a particular domain.

The knowledge-based approach helps to solve the problem of word sense disambiguation for effective information extraction. Word sense disambiguation is essential for information extraction [8], because the same words in different contexts can have different meaning.

Since ontologies explicitly represent knowledge domain semantics (terms in the domain and relations among them), they can be effectively used in solving information extraction problems, word sense disambiguation in particular.

2. Implementation of the proposed semantic patterns approach

The proposed approach has been successfully realized in BuzzTalk portal [9] for subject domains recognition, opinion mining, mood state detection, event extraction and economic activities detection.

BuzzTalk is offered to companies as a SaaS model (Software as a Service) and it will answer questions like:

Who are relevant influencers in my industry?

What are people writing about my brand, product, organization or CEO?

What are important trends in my industry?

What are the big events inside my industry sector?

Where are my influencers/customers/clients located?

When and where are people discussing my brand?

2.1. Subject Domains Recognition in BuzzTalk

A subject domain is recognized on the basis of a particular set of noun and verb phrases unambiguously describing the domain. For solving the problem of disambiguation special filters, based on the contextual environment (on the level of phrases and the whole text), are introduced.

Subject domains and their concepts are organized hierarchically to state “part-of”, “is a kind of” relations. For example, the subject domain “Medical Care” includes:

Medical Care:

- medical service;
- medical staff;
- disease prevention;
- disease treatment;
- medications.

Based on the proposed approach, an object of the particular class of interest may have its own specific sets of sub-classes. For example, in the automobile domain they can be: engine, transmission, suspension, etc.

Associative relationships, which relate concepts across the tree structure, are also taken into consideration: 1) nominative relationships describing the names of concepts; 2) locative relationships describing the location of one concept with respect to another; 3) associative relationships that represent, for example, the functions, processes a concept has or is involved in; 4) cause-effect relationships.

2.2. Opinion mining in BuzzTalk

The proposed ontology-based approach for semantic patterns actualization was realized in the developed knowledge base, which contains opinion words expressing:

- 1) personal emotional state (e.g. happy, delighted, proud, sad, angry, horrified);
- 2) appreciation (e.g. flexible, efficient, stable, reduced, ideal, backward, poor, highest);
- 3) judgement (e.g. active, decisive, caring, dedicated, intelligent, negligent, evil)

While “judgement” evaluates human behaviors, “appreciation” typically deals with natural objects, manufactured objects, as well as more abstract entities, such as plans and policies. Humans may also be evaluated by means of “appreciation”, rather than “judgement”, when viewed more as entities than as participants, e.g. *lovely medical staff*.

Opinion words can be expressed by: an adjective (*brilliant, reliable*); a verb (*like, love, hate, blame*); a noun (*garbage, triumph, catastrophe*); a phrase (*easy to use, simple to use*). Adjectives derive almost all disambiguating information from the nouns they modify, and nouns are best disambiguated by directly adjacent adjectives or nouns.

Information about the force of evaluation (low, high, the highest) and orientation (positive/negative) is also included in the knowledge base. For example, *safe* (low force, positive orientation), *safer* (high force, positive orientation), *the safest* (the highest force, positive orientation), *unsafe* (low force, negative orientation)

In the knowledge base opinion words go together with their accompanying words, thus forming “opinion collocations” (e.g. *deep depression, deep devotion, warm greetings, discuss calmly, beautifully furnished*). By an “opinion collocation” we understand a combination of an opinion word and accompanying words, which commonly occur together in an opinion-oriented text. The use of opinion collocations is a way to solve the problem of opinion word sense disambiguation (e.g. *well-balanced political leader* and *well-balanced wheel*) and to exclude words that do not relate to opinions (cf. *attractive idea* and *attractive energy*).

We assume that the number of opinion collocations, which can be listed in a knowledge base, is fixed.

The use of opinion collocations within the ontology-based approach opens a possibility to assign names of knowledge domains to them, because opinion collocations are generally domain specific. For example, *helpful medical staff* (“health care”), *helpful hotel reception staff* (“travel-hotel”), *stable economy* (“economics”), *well-balanced politician* (“politics”). More than one knowledge domain may be assigned to an opinion collocation, e.g. *fast service* (“economics-company”, “travel-hotel”).

Processing of the extracted opinion collocations is carried out in their contextual environment. The developed algorithm checks for the presence of modifiers that can change the force of evaluation and orientation indicated in the knowledge base.

Let’s consider the following example: *reliable company*. This opinion collocation has the following information in the knowledge base: “low force” of evaluation and “positive orientation”.

e.g. Honda Motor Company is a *reliable company*.

The evaluation force is changed to “higher force” in *more reliable company*.

e.g. ESP is a *more reliable company* in terms of quality.

The algorithm changes the force of evaluation to “the highest force” when processing the opinion collocations *very reliable company, the most reliable company, extremely reliable company*.

e.g. Apple Inc. is a *very reliable company*.

The orientation is changed to the opposite (“negative orientation”) in the following examples: *unreliable company, not reliable company* (“low force”), *the most unreliable company* (“the highest force”).

e.g. With regards to security, Websense may be *the most unreliable company*.

The opinion collocation *not reliable enough company* has positive orientation, but the “low force” is weakened.

e.g. Obviously because InPlant is *not a reliable enough company*.

“The highest force” of evaluation is weakened in *not a very reliable company* (“positive orientation”).

e.g. MedZilla is *not a very reliable company* in terms of loyalty to the sales force.

The developed knowledge base also provides additional information about quality characteristics and relationships for different objects on which an opinion is expressed (e.g. *software product* evaluation

includes: usability, reliability, efficiency, reusability, maintainability, portability, testability; *travel-hotel* evaluation includes: value, rooms, location, cleanliness, check in/front desk, service).

For example: “The *location* of the Golden Well hotel is *excellent*. The hotel is *beautifully furnished* without being overdone. *Check-in* was *fast* and *easy*. The room was *fabulous*, and the *breakfasts* *amazing*. The *bed* was *comfortable* and the *bathroom* was a *pleasure*. *Friendly* and *attentive* staff.”

2.3. Mood state detection in BuzzTalk

A valuable addition to opinion mining is detection of individual/public mood states. The relationship between mood states and different human activities has proven a popular area of research [10].

BuzzTalk mood detection uses the classification of the widely-accepted “Profile of Mood States” (POMS), originally developed by McNair, Lorr and Droppleman [11].

In BuzzTalk, mood state detection is based on: 1) mood indicators (e.g. “I feel”, “makes me feel”, etc.); 2) mood words (e.g. anger, fury, horrified, tired, taken aback, depressed, optimistic); 3) special contextual rules to avoid ambiguity. BuzzTalk automatically recognizes the following mood states: “Anger”, “Tension”, “Fatigue”, “Confusion”, “Depression”, “Vigor”.

Mood state detection alongside with opinion mining can give answers to where we are now and where will be in future.

2.4. Event extraction in BuzzTalk

The developed algorithm performs real-time extraction of 35 events, the recognition of which is vitally important for decision making in different spheres of business, legal and social activities. The events include: "Environmental Issues", "Natural Disaster", "Health Issues", "Energy Issues", "Merger & Acquisition", "Company Reorganization", "Competitive Product/Company", "Money Market", "Product Release", "Bankruptcy", "Bribery & Corruption", "Fraud & Forgery", "Treason", "Hijacking", "Illegal Business", "Sex Abuse", "Conflict", "Conflict Resolution", "Social Life", etc.

2.5. Economic activities detection in BuzzTalk

BuzzTalk detects 233 economic activities from texts in a natural language. The economic activities cover all major activities represented in NACE classification (Statistical Classification of Economic Activities in the European Community), which is similar to the Standard Industrial Classification and North American Industry Classification System. Each of the detected economic activities has a corresponding NACE code.

Conclusion

The proposed semantic patterns approach has been successfully realized in BuzzTalk portal for opinion mining, mood state detection, event extraction and economic activities detection. The approach ensures high accuracy, flexibility for customization and future diverse applications for information extraction.

Semantic word collocations are a major factor in the development of a wide variety of applications including information extraction and information management (retrieval, clustering, categorization, etc.).

Implementation results show that the proposed knowledge-based approach (with statistical methods involved to prevent unwanted results) is correct and justified and the technique is highly effective.

The proposed approach will be enhanced with reasoning modules and sophisticated algorithms to extract more meaningful information from texts in a natural language.

References

1. Moens, M. Information Extraction: Algorithms and Prospects in a Retrieval Context / M. Moens. – Springer, 2006. – 246 p.
2. Baeza-Yates, R. Modern Information Retrieval: The Concepts and Technology behind Search / R. Baeza-Yates, B. Ribeiro-Neto. – Addison-Wesley Professional, 2011. – 944 p.
3. Buettcher, S. Information Retrieval: Implementing and Evaluating Search Engines / S. Buettcher, C. Clarke, G. Cormack. – MIT Press, 2010. – 632 p.
4. Cruse, D.A. Lexical Semantics / D.A. Cruse. – Cambridge University Press, 1986. – 310 p.

5. Manning, C.D. Foundations of statistical natural language processing / C.D. Manning, H. Schütze. – Cambridge, MA : MIT Press, 1999. – 620 p.
6. An Ontology-Based Approach to Opinion Mining / V. Bilan [et al.] // Proceedings of 10th Intern. Conference PRIP. – Minsk, 2009. – P. 257–259.
7. Fensel, D. Foundations for the Web of Information and Services: A Review of 20 Years of Semantic Web Research / D. Fensel. – Springer, 2011. – 416 p.
8. Eneko, A. Word Sense Disambiguation: Algorithms and Applications / A. Eneko, P. Edmonds. – Dordrecht : Springer, 2006. – 364 p.
9. BuzzTalk [Electronic resource]. – Mode of access : <http://www.buzztalkmonitor.com>. – Date of access : 16.04.2014.
10. Clark, A.V. Mood State and Health / A.V. Clark. – Nova Publishers, 2005. – 213 p.
11. McNair, D.M. Profile of Mood States / D.M. McNair, M. Lorr, L.F. Droppleman // Educational and Industrial Testing Service. – San Diego, California, 1971.

STRUCTURE-FUNCTIONAL APPROACH TO MODULATING FUNCTION TREND ESTIMATION BASED ON TIME-SERIES DATA UNDER UNCERTAIN CONDITIONS

A. Britenkov, A. Dokuchaev
Lobachevsky State University of Nizhny Novgorod, Russia
e-mail: jkd@yandex.ru

The article describes practical issues of experimental signals approximation using orthogonal polynomials on the example of forecast patterns of oscillations of enterprise electrical energy consumption. It is important to use stable and efficient algorithms as well as system approach to make accurate number forecasts based on time-series representations. It is highly effective to use Laguerre, Legendre and Hermite functions of high order for the time-series computer forecasts. Classical orthogonal bases of continuous argument can be used in forecast if consider it to be a special case of extrapolation. The results are compared to auto-regression mode estimation. The use of additional data and structure-functional approach to modulating function is offered to improve forecast accuracy.

Introduction

At the present moment there are two main ways of making forecasts in technical, economic and social spheres. The first one is actual data extrapolation [1] with exact values (e.g. price, currency exchange rate, consumption, demand range, etc.). The second one is a complex of methods based on experts knowledge. However, the best results can be achieved using both ways in complex.

The first method is limited by certain instability of extrapolation function. This leads to forecast instability. The main drawback of the second method is that it is impossible for an expert to consider all the factors affecting the system. These effects can be partially reduced by experts' estimation of the forecasts variants, though the error probability is high as the method is based on subjective component of choice [2].

The complex of both above-mentioned methods makes it possible to form more accurate forecasts of technical open-system functioning. The present article describes mathematical modeling techniques, such as generalized spectral-analytical method [3] combined with the structure-functional approach [4]. The main idea of the forecast modeling described below is to recognize the specific typical patterns of the researched signals or data [5].

1. Methods and approaches

The main objective of the present study is to estimate the trend pattern of the investigated system. To accomplish this, forecast functions with both statistical and fuzzy variables are used.

It is accepted that a one-dimensional timeline is represented by a sequence of number values of some process during the discrete moments of time [6]. As a rule, values are measured at regular intervals. Still, in these methods of time number analysis there is a certain complexity of forecast accuracy. While using these methods it is necessary to put special emphasis on constant component allocation by extrapolation and on estimated values determination [7].

In regression models [1, 6] the discrete representation of the source data is used.

$$s(t) = \{S_0, S_1, \dots, S_i, \dots, S_n\}. \quad (1)$$

Regression analysis is widely used in forecasting. It is especially effective when used in computer calculations. The peculiarity of models in linear regressions is that the dependent variable, R_i is a [linear combination](#) of the parameters which should not necessarily be interdependent. For example, in [simple linear regression](#)

$$s(t) = \sum_{i=1}^p A_i R(t-i) + \varepsilon_i, \quad (2)$$

where $R=R(t)$ – function under investigation and a set of indicators at discrete time period t ; p – number of the nearest previous values of R , ε_i – noise.

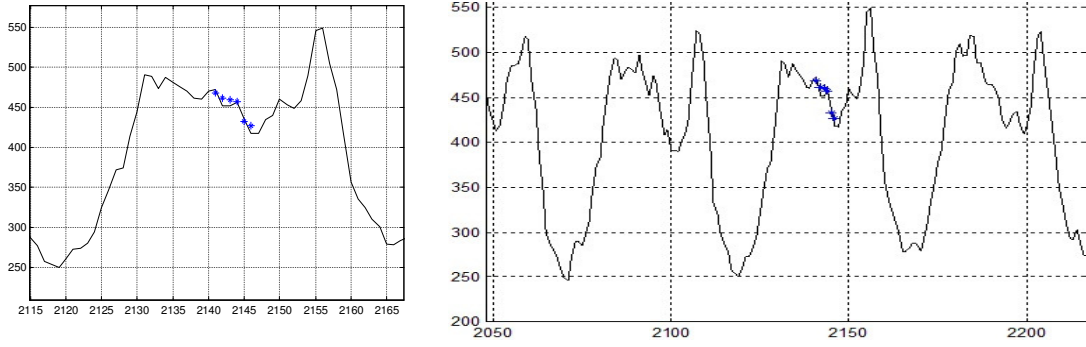


Fig. 1. Energy consumption estimation by autoregressive model ($AR(p), p = 480$). On a plot of the forecast data points are marked in bold. Sample length $N = 100$ points (48 samples per day). Forecast made 6 steps forward (3 h)

2. Time series estimation using orthogonal basis of continuous argument

It seems conclusive to use an enterprise electric energy consumption analysis to illustrate the above mentioned method application.

While constructing energy consumption forecast function (fig. 2) (discrete-time sampling $\Delta t = 30$ min) it is necessary to consider possible noise occurrence. So the forecast function preserves the shape of the experimental curve, but with a reasonable degree of smoothing [8].

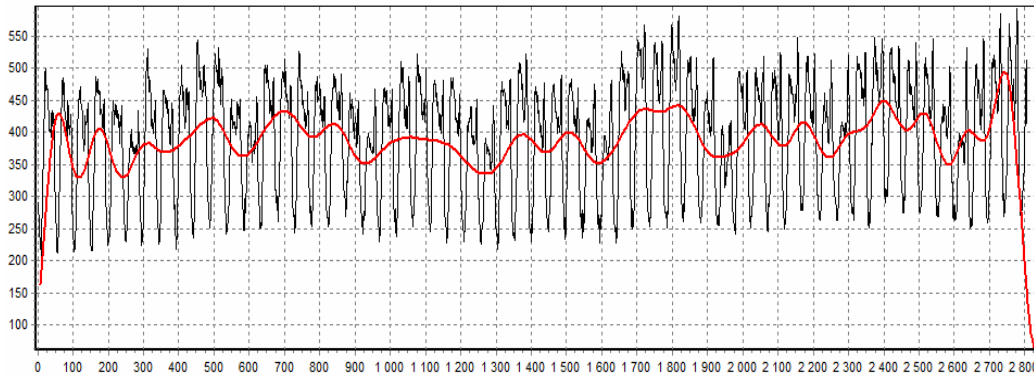


Fig. 2. Approximation of the curve of energy consumption (kWh) enterprises (data marked black orthogonal functions (the red line). Samples taken for 58 days. Horizontal axis - time readout 1 day is equal to 48 readout), vertical axis – power indication values (kWh)

As it can be seen from fig. 2 day and week time intervals can easily be distinguished on the trend line. In order to achieve such distinct results approximating functions [7] are used in this method. Approximation by means of the classical orthogonal functions satisfies the above mentioned criteria. The time sequence $s(t)$ is described by orthogonal series based on modified classical orthogonal polynomials $\{\varphi_i\}$ (which are a special case of the solution of the hyper-geometrical equation [9]) of the classical formula [5]

$$s(t) = \sum_{i=0}^N A_i \varphi_i(t). \quad (3)$$

The coefficients A_i [4] of expansion (3) are determined by a set of linear equations derived by multiplying Eq. (2) by set $\{\varphi_i(t)\}$:

$$A_i = \|\varphi_i\|^{-2} \int_0^T \varphi_i(t) f(t) \rho(t) dt, \quad (4)$$

where $f(t)$ is known data of the researched curve (time series). Natural approximative properties of polynomials (3) reduce the effect of nonessential "outliers" and give information about specific oscillation periods [10].

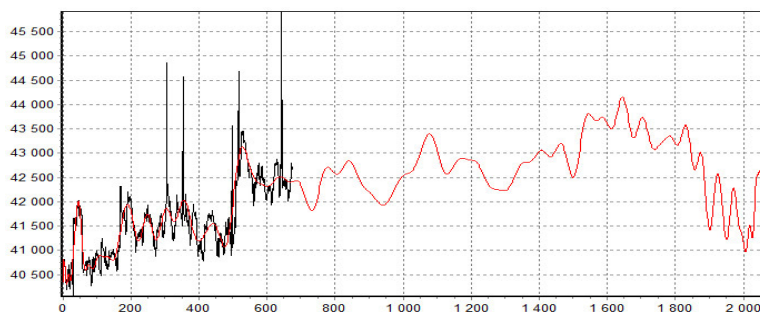


Fig. 3. The enterprise electricity consumption forecast (kW / h) Legendre functions ($P_n(t)$, $n=100$, the number of steps of the forecast is more than 1000, the number of "learning points" $p=675$). Smooth curve - the forecast

The generalized spectral-analytical method is used to forecast the data flow. The forecast problem as a special case of extrapolation and filtering noise comes from the fact that the OLS regression methods used in the forecast are identical to OLS used in Fourier coefficients calculation [5]. It is important to use already developed methods which can use the properties of orthogonal bases on the basis of classical polynomials of continuous argument (e.g. in tasks where trigonometric Fourier series are less effective[11]) to extend the capabilities of the mathematical apparatus of the generalized spectral-analytical method [4]. Physical model of the system comprising N reflecting surfaces for the acoustic wave resembles regression model. Indeed, in such a system, a sound source signal $S(t)$ is transformed to $S^*(t)$. $S^*(t)$ can be represented as the sum of discrete delayed pulses and attenuated after passing through an acoustic medium of the original signal

$$S^*(t) = \alpha_0 S(t) + \sum_{i=1}^N \alpha_i S(t - \tau_i) = \sum_{i=0}^N \alpha_i S(t - \tau_i), \quad (5)$$

where α_i – loss factor; τ_i – time lag one reflection, N – number of reflections which reached the receiver [12]. The possibility to resolve the inverse task of the wave propagation in terms of discrete reflections by applying the generalized spectral-analytical method allows using this technology as a component of regression modeling.

Smoothing function describes the known values of the time series [13] at training time from t_0 to t . Some specific parameters may not be known for this function at time t_0, t_1, \dots, t_p . In this case it is necessary to use modulating function $f_{\text{mod}}(t)$ to construct forecast function taking into account the possible distribution decision of smoothing function parameters based on certain values [14]. The resulting function of the forecast time for this

$$s^*(t_i) = s(t) f_{\text{mod}}(t). \quad (6)$$

Settings of this modulating function can be determined by structure-functional approach or the method of expert estimates or other [2, 6]. In fact, the main goal of the modulating function parameters determination is to find changing points of the trend [1]. This method may be useful for trade and enterprises, as well as for forecasting atmosphere phenomena.

4. Using additional data

Regression analysis can be used to determine cause-and-effect relationship between the independent and dependent variables [1]. The approach to the analysis of the enterprise electric energy consumption with the use of additional data (such as time of day and day of a week on table) gives more accuracy in power consumption estimation for a longer period of time.

Circle representation of consumption time sampling using additional data

Circle representation	Clock	Sample
0	12 p.m.	1
1	6 a.m.	12
0	12 a.m.	24
-1	6 p.m.	36

It can be seen in the two graphs constructed (fig. 4, 5) on the same time-scale how cyclical the consumption of electric energy is. This fact allows using long time scales, usual trigonometric series, but

at short intervals (up to days) combined with a more appropriate description in the form of classical orthogonal functions [8, 13]. Thus it is possible to determine modulating function settings

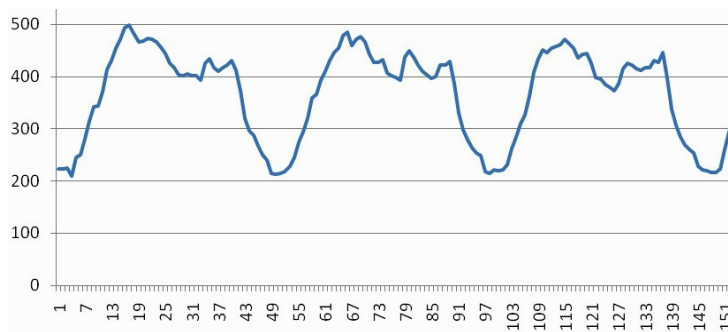


Fig. 4. A pattern of power consumption (kWh) for 3 days

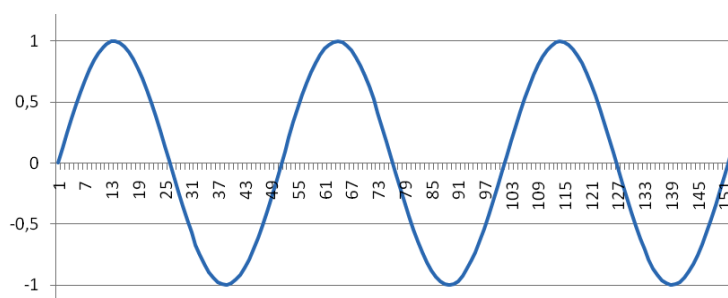


Fig. 5. Daily time in circle representation (one day is 48 samples)

Conclusion

Modulation function application allows considering not only known time series, but also the influence of factors not accounted for in the earlier series of values while making forecasts. Thus it is possible to get a more accurate estimation. However, for each system requires a unique selection of modulation functions that makes it necessary to develop specific procedures for each system. Another problem of the presented idea is that the calculation of modulation functions value may be taken only for the time t , due to the fact that some part of the argument function cannot be estimated and may be added only at the moment of happening.

The study shows that a simple solution using the ratio of the long-range differences gives a good approximation to the solution of the problem, as well as solution based on the optimal analytical approximation [8, 13]. Optimization in this case means getting the type of approximating function, which would reveal the correlation between power consumption and some other factors (seasonal effects, fluctuations of electricity costs, prices, etc). The described approach using system analysis [2] of researched time series (consumption of electric energy at the industrial enterprise, market indicators, etc) according to the facts of temporary fluctuations [14] can be useful for many spheres of life for a long period of time.

References

1. Robert, K.O. Applied Time Series Analysis / K.O. Robert, E. Loren. – N.Y. ; Chichester ; Brisbane ; Toronto : A Wiley-Interscience Publication John Wiley and Sons, 1978. – 428 p.
2. Voinov, B.S. Information technology and systems / B.S. Voinov. – M. : Nauka, 2003. – 653 p.
3. Generalized Spectral Analytical Method for Processing of Information Arrays. Problems of Image Analysis and Recognition / F.F. Dedus [et al.]; ed. F.F. Dedus. – M. : Mashinostroenie, 1999. – 357 p. [in Russian]
4. Zwicky, F. Discovery, Invention, Research through the Morphological Approach / F. Zwicky. – N.Y. : McMillan, 1969. – 276 p.
5. Analytical description of multidimensional signals for solving problems of pattern recognition and image analysis / A.F. Dedus [et al.] // Pattern Recognition and Image Analysis. – 1993. – Vol. 3. – P. 459–469.

6. Ruey, S. Tsay Analysis of Financial Time Series / S. Tsay Ruey // Wiley Series in Probability and Statistics. – N.Y. : Wiley and Sons, 2005. – 576 p.
7. Wiener, N. Extrapolation, interpolation, and smoothing of stationary time series / N. Wiener. – N.Y. : J. Wiley, 1949. – 170 p.
8. Britenkov, A. Forecasting of time series by classical orthogonal functions in regression models / A. Britenkov // Proc. of 8th Open German-Russian workshop "Pattern Recognition and Image Understanding OGRW-8-2011". – N. Novgorod : Nizhny Novgorod Lobachevsky State University. – 2011. – P. 28–30.
9. Kazuhiko, A. Theory of Hypergeometric Functions / A. Kazuhiko, K. Michitake // Springer Monographs in Mathematics Series. – 2011. – Vol. 305. – 317 p.
10. Britenkov, A.K. Stable algorithms of adaptive approximation for acoustic signals description by orthogonal polynomials / A.K. Britenkov, A.N. Pankratov // Physics of Wave Phenomena. – 2004. – Vol. 12 (3). – P. 168–174.
11. Khabibrakhmanv, I.K. The Use of Generalized Laguerre Polynomials in Spectral Methods for Nonlinear Equations / I.K. Khabibrakhmanv, D. Summers // Computers Math. Appl. – 1998. – Vol. 36 (2). – P. 65–70.
12. Britenkov, A.K. Suppression of multiplicative noise in terms of discrete reflections by generalized spectral-analytical method / A.K. Britenkov, A.N. Pankratov // Vestnik of Lobachevsky State University. – Nizhny Novgorod, 2006. – Vol. 1 (4). – P. 50–57. [in Russian]
13. Pankratov, A. Long-Range Monitoring of Fuel Consumption of Car Based on Generalized Spectral-Analytical Method / A. Pankratov // Book Of Proc. "Driver Car Interaction & Interface 2010". – Prague, 2010. – P. 49–50.
14. Shumafov, M.M. Development of the numerical solution algorithm for the inverse task of filtration by modulating functions / M.M. Shumafov, R. Tsey // Vestnik of Adyghe State University. – 2009. – Vol. 4 (1). – P.43–49. [in Russian]

RECONFIGURING DRIVING STYLES BASED ON BAYESIAN NETWORKING PRINCIPLES

G. Dimitrakopoulos

Harokopion University of Athens, Department of Informatics and Telematics, Greece

e-mail: gdimitra@hua.gr

The ceaseless evolution of Information and Communication Technologies (ICT) is reflected on their migration towards the Future Internet (FI) era, which is characterized, among others, by powerful and complex network infrastructures, and innovative applications, services and content. An area of applications that finds prosperous ground in the FI era lies in the world of transportation. The goal of this paper is to introduce an Intelligent Transportation System (ITS) that utilizes (i) the driver's preferences, (ii) information extracted from the vehicle sensors, and (iii) previous knowledge and experience, in proposing adaptations of the vehicle's driving style, in an automated manner. Knowledge is obtained through the exploitation of Bayesian networking concepts and specifically the Naïve-based model. Some indicative simulation results showcase the effectiveness of the proposed system, the advantage of which lies in that the reliability of the knowledge-based selection decisions is higher.

Introduction

Information and Communication Technology (ICT) continue to attract immense research interest [1, 2]. Latest trends refer to the migration of ICT towards the era of the “Future Internet” (FI) [3], which envisages mechanisms that promise easier overcoming of the structural limitations of telecommunication infrastructures and their management systems, so as to further facilitate the design, development and integration of novel services and applications [3, 4]. An application area that finds prosperous ground in the FI era is transportation. Lately, the automotive world is witnessing a trend related to the extensive use of telecommunications systems inside vehicles. The results of such trends are reflected on the term “ITS”, which envisages systems that are either related to road infrastructures, making the infrastructure “intelligent”, or used inside vehicles traveling on road, attributing vehicles with intelligence [1–5]. Sensors and sensor networks play a significant part in this effort, as today they are broadly used in passenger vehicles, for safety, as well as emission control reasons. In the future, vehicles will be capable of offering more extensive navigation assistance, monitoring their own systems and behaviour, reconfiguring their operating parameters and alerting the driver when action is required, through sensor systems that will help drivers to cope with hazardous conditions.

With the vision to build on the aforementioned research approaches, the motivation for the work presented in this paper is the fact that several parameters that affect the selection of the appropriate matches among drivers and driving styles, can be changing with time (in a random manner). Therefore, system that can increase the reliability of the decisions is required. The system should provide the probability that the parameters will achieve certain values, based on specific matches. The learning functionality is influenced by Bayesian networks [6–8]. The structure selected in this paper is based on the Naive-based Model [8]. This model simplifies learning by assuming that features (parameters in our case) are independent given the class (selected match in our case). The conference will provide a forum for scientists and engineers to exchange up-to-date technical knowledge and experience in the field of image analysis, recognition and information processing, define ways of further development of this subject. The conference will focus on both theory and applications.

1. Business case and high level description

A set of drivers that may drive a certain vehicle is assumed (in the case of a family, usually one of them is the most frequent driver), as well as a set of driving styles. The drivers and the driving styles are associated with specific parameters, i.e. (a) context information deriving from measurements obtained from the vehicle's sensors, data on the driver's personal profile parameters, and on the other hand data associated with style related parameters. Last, a set of overarching policies reflects driver/styles preferences, in the form of weights (importance) attributed to the aforementioned parameters.

In general, the manner in which a driver operates the vehicle can change from time to time. This is depicted in a change of the personal profile parameters. Thus, a change in the driving style of the vehicle may be desirable (change of suspension adjustments, gear ratios, speed of vehicle reaction, etc.). The goal of the system is to interact, on behalf of the driver, with all candidate driving styles and find and propose

an optimum match. Communication can be guaranteed through the existence of an, easy to deploy, ICT-based management system (such as the one proposed herein – *a-drive*).

In the light of the above, *a-drive* is shown on fig. 1. It uses as input (i) personal profile parameters, (ii) vehicle sensor measurements and (iii) policies which attribute importance to the parameters through numerical weights. The output of the algorithmic functionality is the optimum matching among drivers and driving styles. The solution method follows a phased approach, consisting of (i) the “robust discovery phase” and (ii) the “decision making phase”. The robust discovery phase aims at maximizing the probabilities that the parameters will reach certain values, through a Bayesian based model, which helps the system obtain knowledge. The decision making phase steps on those probabilities and finds the optimum matching considering also the importance of the parameters.

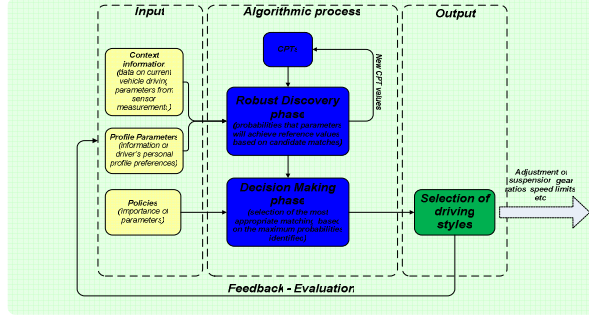


Fig. 1. *A-drive* high-level description

It should be also noted that knowledge acquisition is further enhanced by an evaluation procedure, made by the driver concerning driving styles after the completion of a ride. In this respect, parameters are evaluated, at an integers’ scale from “1” to “10”, in the form of utility volumes [16], with “1” standing for “poor” and “10” standing for “excellent”.

2. Formal description

Input. The focus is on a driver that drives a vehicle equipped with the *a-drive* system. The set of the potential vehicle’s drivers is PD . D is defined for representing the driver. D can take values 1 to $|PD|$. In the same manner, the set of candidate driving styles is denoted as CDS . DS is defined for representing the driving style. DS can take values 1 to $|CDS|$. The set of parameters is denoted as N . Each parameter, j ($j=1, \dots, N$), can refer to a specific aspect, e.g. mean driving speed, age, gender, etc. Finally, the importance of each parameter, j ($j=1, \dots, N$) is indicated by a weight value w_j . In principle, the sum of the w_j weights, over all $j=1, \dots, N$, will be 1. The w_j values can constitute a vector of weights \mathcal{W} . As previously mentioned, the parameter values can be changing with time, in a random manner. Therefore, variable i is defined for representing the driving style. Moreover, variable v_j ($j=1, \dots, N$) depicts the value of the j -th parameter. Each variable V_j is associated with a set of reference values RV_{ij} ($i \in CDS$). Variable v_j can take a value among those in RV_{ij} , when driving style i is considered.

The knowledge that needs to be developed relies on conditional probabilities, which have the form $\Pr[V_j = rv_{ij}^k | DS = i]$, where $rv_{ij}^k \in RV_{ij}$ denotes the k -th reference value for the j -th parameter when driving style i is considered.

Probability density function. The following probability density function can be defined:

$$f(\tilde{x}, i) = \Pr[V_1 = rv_{i1}^{k_1}, \dots, V_N = rv_{iN}^{k_N}, DS = i] = \Pr[DS = i] \cdot \prod_{j=1}^N \Pr[V_j = rv_{ij}^{k_j} | DS = i], \quad (1)$$

where $i \in CDS$, $\tilde{x} \in \mathcal{X}_i$, $rv_{ij}^{k_j} \in RV_{ij}$ ($j=1, \dots, N$), and k_j ($j=1, \dots, N$) are integers.

The $\Pr[DS = i]$ probabilities show the volume of information existing for each driving style i . The sum of the $\Pr[DS = i]$ quantities, over all $i \in CDS$, is 1. The more information there is on a driving style- i the more reliable the knowledge, and therefore, the higher the $f(\tilde{x}, i)$ values.

In general, the values of the $f(\tilde{x}, i)$ function express in an aggregate manner our knowledge on how probable is the achievement of a parameter value indicated in \tilde{x} , by driving style i .

Objective and solution

The objective is to select the most appropriate driving style among those in CDS . To do so, the proposed algorithm follows two phases, as previously mentioned, namely (i) the robust discovery phase and (ii) the decision making phase.

Robust discovery phase. The goal of this process is to identify the most probable parameter values. To do so, the probabilities in the right end of (1) need to be updated. For this purpose, *a-drive* collects evaluations made for the CDS driving styles. The update of the conditional probabilities in relation (1) can take into account the “distance” of the collected evaluation values from the reference values. Let us assume that the most recent evaluation indicates that driving style i can achieve rv_{ij}^{coll} regarding parameter j . Let dif_{ij} be the difference between the maximum and the minimum reference value in RV_{ij} . Then, for each reference value, $rv_{ij}^k \in RV_{ij}$, there can be a correction factor:

$$cor_{ij}^k = 1 - (|rv_{ij}^k - rv_{ij}^{coll}| / dif_{ij}). \quad (2)$$

Since $0 \leq cor_{ij}^k \leq 1$, a value close to one means that the reference and collected values are close, and thus, that the corresponding conditional probability value should be reinforced accordingly. The opposite holds, if cor_{ij}^k is close to zero. The new conditional probabilities are obtained through the following relation:

$$\Pr[V_j = rv_{ij}^k | DS = i]_{new} = nf_{ij} \cdot cor_{ij}^k \cdot \Pr[V_j = rv_{ij}^k | DS = i]_{old}. \quad (3)$$

Parameter nf_{ij} is a normalizing factor for guaranteeing that all the “new” probabilities will sum up to one. Moreover, in order to ensure adaptability to new conditions, the conditional probabilities can be prohibited from exceeding a certain threshold, pr_{max} . In summary, the update strategy includes: (i) collection of parameter reference values (through evaluations and measurements); (ii) computation of the correction factors through relation (2), and of the new probabilities through relation (3); (iii) if a probability exceeds pr_{max} it is set equal to the threshold; (iv) the new normalizing factors are calculated, by forcing the remaining probabilities to sum to $(1 - pr_{max})$, and the new values are computed for the remaining probabilities.

Decision making phase – exploitation of knowledge. The scheme favors the selection of driving styles that have high probability of achieving the most appropriate parameter values (thus living up to the driver expectations). In order to model these aspects an Objective Function (OF) value, OF_i , is defined for each driving style, $i \in CDS$. The computation of the OF values, OF_i , of all driving styles $i \in CDS$, is made through the following relation:

$$OF_i = \sum_j \left\{ \max(\Pr[V_j = rv_{ij}^k | DS = i]) \right\} \cdot w_j, \quad (4)$$

where $i \in CDS$, ($j = 1, \dots, N$) and $rv_{ij}^k \in RV_{ij}$ denotes the k -th reference value for the j -the parameter when driving style i is considered. The driving style with the highest OF_i value should be selected based on the knowledge obtained from the aforementioned process. In particular, each driving style

corresponds to a specific combination of (a) suspension adjustment, (b) gear ratios, (c) speed limits and (d) steering wheel reciprocation. Finally, the decision is implemented.

3. Results

The scenario used aims at showcasing the gradual development of knowledge and the impact of the continuous change of a driver (who gradually drives more smoothly) on the decision making process. Last, 3 different driving styles are assumed, namely comfort, normal and sport. As mentioned above, each driving style corresponds to a specific combination of (a) suspension adjustment, (b) gear ratios, (c) speed limits and (d) steering wheel reciprocation. In particular, the scenario assumes that the driver changes his driving behavior, i.e. from a more aggressive one towards a more conservative one. The parameter values (obtained either through sensors or inserted by the driver during the evaluation process) are omitted for brevity reasons. For facilitating the process, it is assumed that 15 computations are split in 3 phases (each one lasting for 5 computations). The second driving style exhibits a better performance in each subsequent phase, implying that it is more suitable.

Fig. 2, a depicts the conditional probabilities of parameter “economy”, likely to be achieved by the second driving style, split in 3 phases. In the first phase which lasts for 5 computations), the conditional probability $\Pr[V_{economy} = 7 | DS = 2]$ appears to be the prevalent one. Then, in the second phase (which lasts for computations 6-10), again $\Pr[V_{economy} = 7 | DS = 2]$ is the highest one.

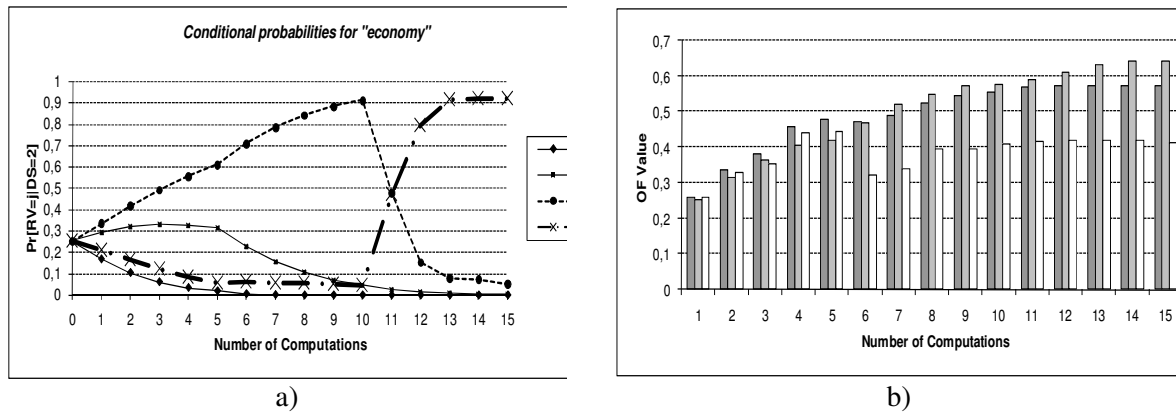


Fig. 2. Conditional probabilities of parameter “economy” of the 2nd driving style in the 3 phases (the driver is assumed to become more conservative) (a); OF values for the 3 driving styles (b)

However, a slight increase in the values of $\Pr[V_{economy} = 10 | DS = 2]$ is observed, with a parallel diminishment of the rest probabilities. Finally, in the third phase (computations 11-15), the most likely reference value to be achieved is 10 and thus $\Pr[V_{economy} = 10 | DS = 2]$ gradually becomes the dominant one. There is naturally a point (computation 11) where a false decision may be taken. However, the system quickly “recovers” and thus the small amount of time consumed for knowledge development is a desirable property. It is the time required in order to increase the reliability levels regarding the new capabilities of the second driving style, in terms of achieving a suitability level with regards to the driver desires/behavior. In this time period the driving style exhibits a “good” behavior. In case the behavior is unstable, the improvement will be considered temporary. The different conditional probabilities will be at low levels, so they will not indicate a clear advantage for any driving style. In any case, however, the amount of time required for the development of knowledge is not large, therefore enabling fast adaptations. The OF values of the 2nd driving style reach the highest possible values after around twelve steps on average. These are depicted on fig. 2, b.

When possible, the 2nd driving style becomes more appropriate in six steps on average. In general, a small computational effort is required for acquiring the knowledge. This is highly desirable, as it can catch improvements in the behavior of a driver even at a non-permanent basis. The number of steps is not high, and therefore, fast adaptations are possible.

Conclusions

This paper has presented an ITS based on Bayesian networking principles, namely *a-drive*, targeted at exploiting knowledge and experience from past interactions, in dynamically proposing the most appropriate driving style for a driver, whilst driving the vehicle. Results show that *a-drive* can (i) adapt to parameter changes fast and successfully and (ii) propose the most suitable driving style whilst driving a vehicle based on knowledge, experience and enhanced decision-making.

This work could be extended by developing further machine learning techniques that could create collective knowledge that would be exploited by *a-drive* more efficiently in reaching the appropriate decisions. Additionally, what could also be investigated is the potential to change the importance (weights) attributed to the parameters during the robust discovery phase and then test the system's response. Part of our future activities shall be also devoted to the integration of the concept of in-vehicle intelligence in larger management functionality for ITS that could exploit several novel concepts, such as issuing directives to the drivers in tackling emergency situations, amending traffic lights and taking other useful decisions during a vehicle's ride.

References

1. Hasselbring, W. Towards trustworthy software systems / W. Hasselbring // IEEE Computer. – 2006. – Vol. 29, no. 4.
2. Project End-to-End Efficiency (E3) [Electronic resource]. – Mode of access : www.ict-e3.eu. – Date of access : 16.04.2014.
3. European Future Internet Initiative (EFII) [Electronic resource]. – Mode of access : <http://initiative.future-internet.eu>. – Date of access : 16.04.2014.
4. Kephart, J. The vision of autonomic computing / J. Kephart, D. Chess // IEEE Computer. – 2003. – Vol. 36, no.1. – P. 41–50.
5. Poole, R. Virtual Exclusive Busways: Improving Urban Transit While Relieving Congestion / R. Poole, T. Balaker // Policy Study 337, Reason Foundation. – September 2005.
6. Russell, S.J. Artificial Intelligence: A Modern Approach / S.J. Russell, P. Norvig. – New Jersey : Prentice-Hall, 2002.
7. Neapolitan, R.E. Learning Bayesian drivers / R.E. Neapolitan. – New Jersey : Prentice-Hall, 2002.
8. Mas-Colell, A. Microeconomics / A. Mas-Colell. – London : Oxford University Press, 1995.

FRACTAL LEARNING OF FAST ORTHOGONAL TWO-DIMENSIONAL TRANSFORMATIONS

A.Yu. Dorogov

St. Petersburg State University of Electrical Engineering "LETI", Russia

e-mail: vaksa2006@yandex.ru

A new method of learning fast two-dimensional orthogonal transformations is considered. Tunable orthogonal transformations are regarded as special neural networks. The learning takes a finite number of steps. The learning algorithm does not have the error feedback and is absolutely stable. The method is based on fractal filtering of signals and images. Linguistic models are used to determine the topology and structure of fast transformations. Examples are given.

Introduction

Signal recognition and classification often require preprocessing procedures that help remove redundancy and select informative features. Use of orthogonal transformations for this purpose allows input information to be presented as independent spectral components. It is known that the greatest reduction of data redundancy is ensured by using Karhunen-Loeve orthogonal transformation formed by the signal covariance matrix eigenvectors. However, using this transformation involves a lot of computations. For this reason Karhunen-Loeve method is not used in processing of data of large dimensionality. However, if only one, most important eigenvector (the principal component) is used in Karhunen-Loeve transformation, the computation load decreases significantly. The orthogonal transformation tuned to one principal component belongs to the class of adapted transformations.

In 1970s Andrews and Caspari [1] were the first who suggested the idea of generalized orthogonal transformation. First learning algorithms for this kind of transformations were developed by A.I. Solodovnikov and his colleagues [2]. This class of transformations was also known as tunable fast transformations at that time. Introducing activation functions and offsets turns a fast transformation into a fast neural network [3] which can be trained by gradient algorithms like for example the error back propagation algorithm. These algorithms can also be used to train linear tunable fast transformations (given bounding nonlinearity in the learning circuit). The drawback of gradient algorithms consists in potentially unstable learning procedures and the presence of dead ends and hanging up at local minima. Structural properties of fast transformations allow specific learning methods free from the above-mentioned drawbacks. Considered below, the method of training adapted transformations is based on the possibility of multiplicative decomposition of fast-transformation matrix elements proved by Good [4] (1958). Methods of fractal finite-interval filtration of signals [5, 6] are used for multiplicative decomposition.

1. System models of two-dimensional fast transformations

Orthogonal transformations with fast execution algorithm are usually used in image processing. The aim of the processing usually involves filtration and compression of an image.

Let us denote a $N_y \times N_x$ image matrix as $F(U_y, U_x)$. When we subject an image to linear transformation $H(U_y, U_x; V_y, V_x)$, we get an array of $M_y \times M_x$ coefficients. The two-dimensional transformation complies with the following rule:

$$S(V_y, V_x) = \sum_{U_y=0}^{N_y-1} \sum_{U_x=0}^{N_x-1} F(U_y, U_x) H(U_y, U_x; V_y, V_x). \quad (1)$$

A two-dimensional transformation is called an orthogonal (unitary, to be exact) transformation if the following condition is met:

$$\sum_{U_y=0}^{N_y-1} \sum_{U_x=0}^{N_x-1} H(U_y, U_x; V_y, V_x) \bar{H}(U_y, U_x; V'_y, V'_x) = \begin{cases} 1 & \text{if } V_y = V'_y \text{ and } V_x = V'_x \\ 0 & \text{if } V_y \neq V'_y \text{ or } V_x \neq V'_x, \end{cases} \quad (2)$$

where the over line sign marks the complex-conjugate transformation. For an orthogonal transformation we have $N_y = M_y$, $N_x = M_x$. The necessary condition of existence of a fast algorithm is the possibility

of multiplicative decomposition of the values of both dimensions of an image into the same number of multiplicands:

$$\begin{aligned} N_y &= p_0^y p_1^y \mathbf{K} p_{n-1}^y, \\ N_x &= p_0^x p_1^x \mathbf{K} p_{n-1}^x. \end{aligned}$$

Here indices x, y indicate to which coordinate axis the original image belongs to. The above condition is not a very strict limitation because some multiplicands can have unit values. Nevertheless, the greater the number of non-unit multiplicands, the higher the computation efficiency of a fast algorithm. Using multiplicands N_y and N_x let us express the coordinates of image points in a positional numeration of compound bases:

$$\begin{aligned} U_y &= \langle u_{n-1}^y u_{n-2}^y \mathbf{L} u_1^y u_0^y \rangle, \\ U_x &= \langle u_{n-1}^x u_{n-2}^x \mathbf{L} u_1^x u_0^x \rangle. \end{aligned} \quad (3)$$

where weight of the m -th digit is determined as $p_{m-1}^* p_{m-2}^* \mathbf{K} p_1^* p_0^*$, and u_m^* is a digit variable taking values $[0, p_m^* - 1]$ (the asterisk replaces indices x and y here). Similarly we can express the coordinates of spectral coefficients in plane $[V_y, V_x]$:

$$\begin{aligned} V_y &= \langle v_{n-1}^y v_{n-2}^y \mathbf{L} v_1^y v_0^y \rangle, \\ V_x &= \langle v_{n-1}^x v_{n-2}^x \mathbf{L} v_1^x v_0^x \rangle. \end{aligned}$$

The algorithm of a fast transformation is usually represented as a graph of different topologies. In the case of Tukey-Cooley topology with time down sampling, the graph can be described as a linguistic sentence [6]:

$$\left[\langle u_{n-1}^* u_{n-2}^* \mathbf{K} u_1^* u_0^* \rangle \langle u_{n-1}^* u_{n-2}^* \mathbf{K} u_1^* v_0^* \rangle \mathbf{K} \langle u_{n-1}^* u_{n-2}^* \mathbf{K} u_{m+1}^* u_m^* v_{m-1}^* v_{m-2}^* \mathbf{K} v_0^* \rangle \mathbf{K} \langle v_{n-1}^* v_{n-2}^* \mathbf{K} v_1^* v_0^* \rangle \right].$$

The first and the last words of the sentence correspond to the coordinates of values in the spatial and spectral areas. The number of layers in the sentence is $n+1$. The words in the middle determine the coordinates U_y^m, U_x^m and V_y^m, V_x^m of points of an input image in the inner layers of the fast algorithm. If the algorithm has a regular compact topology, the following condition is true:

$$U_y^{m+1} = V_y^m, \quad U_x^{m+1} = V_x^m \quad (4)$$

In the general case the topologies for x and y axes can differ. The graph of a fast algorithm in layer m holds base operations $W_{i_x^m, i_y^m}^m(u_m^y u_m^x; v_m^y v_m^x)$ which are four-dimensional matrices of dimensionality $[p_m^y, p_m^x; p_m^y, p_m^x]$ (transformation kernels). The relation between base operations is determined by the fast transformation structural model. For the given topology the graph of the structural model is described by the linguistic sentence:

$$\left[\langle u_{n-1}^* u_{n-2}^* \mathbf{K} u_1^* \rangle \langle u_{n-1}^* u_{n-2}^* \mathbf{K} u_2^* v_0^* \rangle \mathbf{K} \langle u_{n-1}^* u_{n-2}^* \mathbf{K} u_{m+1}^* v_{m-1}^* v_{m-2}^* \mathbf{K} v_0^* \rangle \mathbf{K} \langle v_{n-2}^* v_{n-3}^* \mathbf{K} v_1^* v_0^* \rangle \right].$$

Each word in the sentence defines the number of base operation i_*^m in layer m . The number of words in the sentences is n .

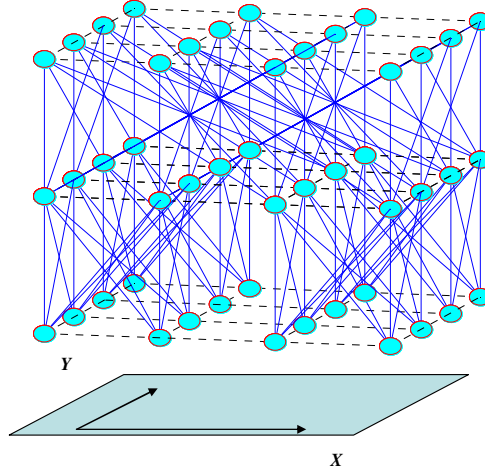


Fig. 1. The structural model of a two-dimensional fast transformation

Fig. 1 shows the structural model of a two-dimensional transformation for an 8×8 image. The input image goes to the bottom layer, while spectral coefficients are produced in the top layer. Base operations relate to the model nodes. A kernel in layer m makes the two-dimensional transformation of a $p_m^y \times p_m^x$ space block:

$$S^m(V_y^m, V_x^m) = \sum_{u_m^y} \sum_{u_m^x} F^m(U_y^m, U_x^m) W_{i_x^m, i_y^m}^m(u_m^y u_m^x; v_m^y v_m^x) \quad (5)$$

Correspondences $U_*^m \leftrightarrow (i_*^m, u_*^m)$ are determined one-one by the linguistic sentences of the topological and structural models.

2. Multiplicative decompositions of matrix elements of the two-dimensional fast transformation

Setting specific values for all digit variables u_m^*, v_m^* (where m takes 0, 1, ..., $n-1$ successively) defines a particular path in the topological graph between a couple of elements from the start and finish layers. It follows from the single-valuedness of the digit representation of numbers that this path is unique for each pair combination of space points of the input and output layers. This fact allows a convenient analytic expression connecting the fast transformation matrix elements with kernel elements. From (1) it follows:

$$H(U_y, U_x; V_y, V_x) = \frac{\partial S(V_y, V_x)}{\partial F(U_y, U_x)} \quad (6)$$

Differentiating (6) as a compound function, we get:

$$H(U_y, U_x; V_y, V_x) = \frac{\partial S^{n-1}}{\partial F^{n-1}} \frac{\partial F^{n-1}}{\partial S^{n-2}} \frac{\partial S^{n-2}}{\partial F^{n-2}} \mathbf{L} \frac{\partial F^1}{\partial S^0} \frac{\partial S^0}{\partial F^0}.$$

From condition (4) it follows that for all m we have $\frac{\partial F^m}{\partial S^{m-1}} = 1$, and from (5) it follows that

$\frac{\partial S^m}{\partial F^m} = W_{i_x^m, i_y^m}^m(u_m^y u_m^x; v_m^y v_m^x)$. So we find that each element of the four-dimensional transformation matrix H is expressed with the help of kernel elements as the product:

$$H(U_y, U_x; V_y, V_x) = W_{i_x^{n-1}, i_y^{n-1}}^{n-1}(u_{n-1}^y u_{n-1}^x; v_{n-1}^y v_{n-1}^x) W_{i_x^{n-2}, i_y^{n-2}}^{n-2}(u_{n-2}^y u_{n-2}^x; v_{n-2}^y v_{n-2}^x) \mathbf{K} W_{i_x^0, i_y^0}^0(u_0^y u_0^x; v_0^y v_0^x), \quad (7)$$

where the digit expressions of kernels indices of layer m for the given topology has the form:

$$\begin{aligned} i_x^m &= \langle u_{n-1}^x u_{n-2}^x \mathbf{K} u_{m+1}^x v_{m-1}^x v_{m-2}^x \mathbf{K} v_0^x \rangle, \\ i_y^m &= \langle u_{n-1}^y u_{n-2}^y \mathbf{K} u_{m+1}^y v_{m-1}^y v_{m-2}^y \mathbf{K} v_0^y \rangle. \end{aligned} \quad (8)$$

By putting expression (7) in the orthogonality condition (2), we find that it will be fulfilled when all kernels are orthogonal, i.e. for any m, i_x^m, i_y^m the following takes place:

$$\sum_{u_m^y} \sum_{u_m^x} W_{i_x^m, i_y^m}^m (u_m^y u_m^x; v_m^y v_m^x) \bar{W}_{i_x^m, i_y^m}^m (u_m^y u_m^x; v_m^y v_m^x) = \begin{cases} 1 & \text{if } v_m^y = \bar{v}_m^y \text{ and } v_m^x = \bar{v}_m^x \\ 0 & \text{if } v_m^y \neq \bar{v}_m^y \text{ or } v_m^x \neq \bar{v}_m^x \end{cases}$$

3. Fractal learning of two-dimensional fast transformations

The learning algorithm for two-dimensional fast transformations uses the ideas of fractal filtration which were discussed in [6]. In the case of two dimensions the fractal filtration is multiple-scale image processing which successively compresses an image down to a single point. The flowchart of the fractal filtration can be pictured as a pyramid shown in Figure 2. The pyramid base is an original image whose $F(U_y, U_x)$ has arguments U_y, U_x are expressed in a positional numeration (see (3)). Let us fix all digits except for two low order digits u_0^y and u_0^x in this positional representation.

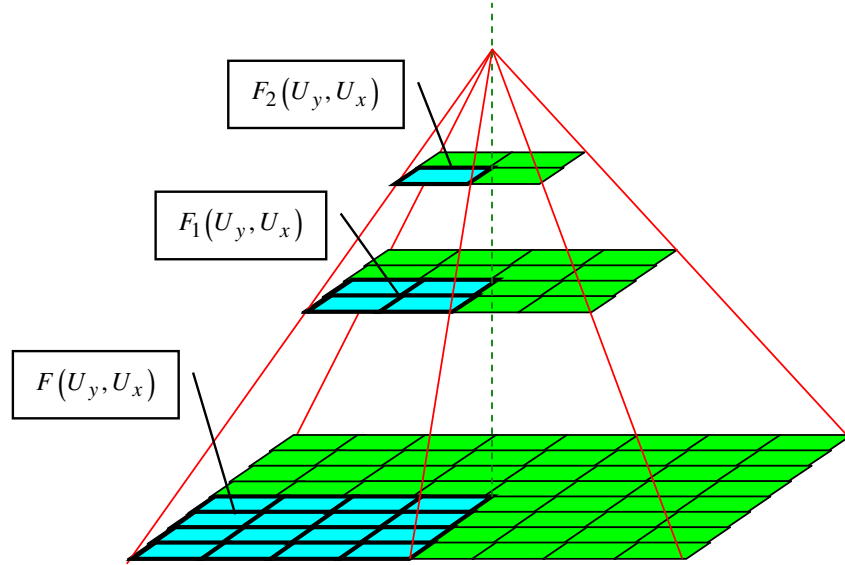


Fig. 2. The diagram of fractal filtration of an image

If we make these digits take all possible values, we get a two-dimensional sampling of size $p_0^y \times p_0^x$. We regard the fractal filter as an arbitrary functional Φ defined over this sampling. Formally it can be written in the form:

$$F_1(\langle u_{n-1}^y u_{n-2}^y \mathbf{L} u_1^y \rangle, \langle u_{n-1}^x u_{n-2}^x \mathbf{L} u_1^x \rangle) = \Phi_{(u_0^y, u_0^x)} \left[F(\langle u_{n-1}^y u_{n-2}^y \mathbf{L} u_1^y u_0^y \rangle, \langle u_{n-1}^x u_{n-2}^x \mathbf{L} u_1^x u_0^x \rangle) \right].$$

It is obvious that image F_1 will be scaled down comparing the size of the original image. The functional may be, for example, the rule of computing the mean of the sampling or its median. The original image can be formally represented as the product:

$$F(\langle u_{n-1}^y u_{n-2}^y \mathbf{L} u_1^y u_0^y \rangle, \langle u_{n-1}^x u_{n-2}^x \mathbf{L} u_1^x u_0^x \rangle) = F_1(\langle u_{n-1}^y u_{n-2}^y \mathbf{L} u_1^y \rangle, \langle u_{n-1}^x u_{n-2}^x \mathbf{L} u_1^x \rangle) f_{j_0^y, j_0^x}(u_0^y, u_0^x),$$
 where $f_{j_0^y, j_0^x}(u_0^y, u_0^x)$ is a set of two-dimensional function-multiplicands which depend on digit variables u_0^y and u_0^x , and indices j_0^y, j_0^x selects the two-dimensional function from the set. The values of the indices

are set equal to the values of the arguments of image F_1 , i.e. $j_0^y = \langle u_{n-1}^y u_{n-2}^y L u_1^y \rangle$ and $j_0^x = \langle u_{n-1}^x u_{n-2}^x L u_1^x \rangle$. In order to get function-multiplicands, it is enough to scalarly divide image F by image F_1 , varying all digit variables in the process. Image F_1 in turn can also be represented as the product of image F_2 and multiplicands from set $f_{j_1^y j_1^x}(u_1^y, u_1^x)$. Repeating the operations of fractal filtration and decomposition over and over again, we will reach the peak of the image pyramid and get multiplicative decomposition of images:

$$F(\langle u_{n-1}^y u_{n-2}^y L u_1^y u_0^y \rangle, \langle u_{n-1}^x u_{n-2}^x L u_1^x u_0^x \rangle) = f_{j_{n-1}^y j_{n-1}^x}(u_{n-1}^y, u_{n-1}^x) f_{j_{n-2}^y j_{n-2}^x}(u_{n-2}^y, u_{n-2}^x) \dots K f_{j_1^y j_1^x}(u_1^y, u_1^x) f_{j_0^y j_0^x}(u_0^y, u_0^x), \quad (9)$$

where $j_m^y = \langle u_{n-1}^y u_{n-2}^y L u_{m+1}^y \rangle$ and $j_m^x = \langle u_{n-1}^x u_{n-2}^x L u_{m+1}^x \rangle$. If we compare the decomposition of an image with the fast-transformation decomposition (7), we will see that they are similar. Moreover, the set of indices of kernels covers the set of indices of function-multiplicands. This fact allows the conclusions that the fast transformation will be accommodated to the image when the transformation kernels are accommodated to decomposition functions (9). In fact, for each two-dimensional function-multiplicand it is necessary to complete orthogonal functions to the full set of basic functions of the orthogonal kernel. It is simple to do by using Gram-Schmidt orthogonalization procedure. The sets of function-multiplicands with deficiency of kernels can be supplemented by themselves. The choice of complementary functions does not have effect on the accommodation condition of the fast transformation, but change the form of other basic functions of the transformation.

Conclusion

We have considered the algorithms of learning orthogonal neural networks that have no closed loop of error-based weight correction. Due to this property the algorithms are always stable and stop running in a finite number of steps, the number of learning steps is proportional to the logarithm of the transformation dimensionality. The precision of learning is fundamentally determined by the processor precision. The learning algorithms for non-orthogonal transformations do not change radically except that there is no Gram-Schmidt orthogonalization procedure.

References

1. Andrews, H.C. A General Techniques for Spectral Analysis / H.C. Andrews, K.L. Caspari // IEEE Tr. Comput. – 1970. – Vol. C19, no. 1. – P. 16–25.
2. Солодовников, А.И. Основы теории и методы спектральной обработки информации / А.И. Солодовников, А.М. Спиваковский. – Ленинград, 1986. – 272 с.
3. Дорогов, А.Ю. Быстрые нейронные сети: проектирование, настройка, приложение [Электронный ресурс] / А.Ю. Дорогов. – Режим доступа : <http://bookfi.org/book/805420>.
4. Good, I.J. The Interaction Algorithm and Practical Fourier Analysis / I.J. Good // J. Royal Statistic. Soc., Ser. B. – 1958. – Vol. 20, no. 2. – P. 361–372.
5. Дорогов, А.Ю. Нейросетевое моделирование регулярных фракталов / А.Ю. Дорогов, М.Ю. Шестопалов // Нейрокомпьютеры: разработка и применение. – 2007. – № 6.
6. Dorogov, A.Yu. Fractal Learning of Fast Orthogonal Neural Networks / A.Yu. Dorogov // Optical Memory and Neural Networks (Information Optics). – 2012. – Vol. 21, no. 2. – P. 105–118.

MEANLESS k -MEANS AS k -MEANLESS CLUSTERING WITH THE BI-PARTIAL APPROACH

S. Dvoenko
Tula State University, Russia
e-mail: dsd@tsu.tula.ru

In a featureless case a set of objects is represented only by results of pairwise mutual comparisons in the form of a distance or similarity matrix. New versions of the well-known k -means algorithm are proposed, where it doesn't need to calculate cluster centers at all. The k -meanless versions of the k -means algorithm make permutations on the comparison matrix and give the same result for both featureless and feature-based cases. According to the bi-partial approach a two-part objective function combines intracluster distances with intercluster similarity and needs to be minimized or combines intracluster similarities with intercluster distance and needs to be maximized.

Introduction

In modern data analysis and data mining, it is the usual way to represent data about a phenomenon under investigation as a matrix of pairwise comparisons of analyzed set members. Such comparisons can be nonnegative values of dissimilarity or similarity of set members. For example, similarity of amino-acid sequences is evaluated by mutual alignment of polymeric chains of protein molecules and in this case, the feature-based representation can't be used at all [1].

A positively definite similarity matrix can be used as a matrix of pairwise scalar products of objects in some metric (Euclidean, as a rule) space with dimensionality not more than a set cardinal number. This matrix of scalar products can be transformed in a distance one and vice versa. As a result, the dissimilarity matrix can be used as the distance matrix in the same space. If set members are features, then the matrix of weighted scalar products (correlations) of them can be used too.

The well-known k -means algorithm is very popular and is simple to understand its results. To get the mathematically correct clustering in featureless case it needs to develop new versions of this algorithm (and other clustering and recognition ones). The basic idea of k -means is the unbiased partitioning. According to it, each cluster Ω_k , $k=1, \dots, K$ is represented by its "representative" object $\tilde{\mathbf{x}}_k$, and the center of the cluster is represented by its "mean" object $\bar{\mathbf{x}}_k$. All cluster centers and representatives from same clusters are the same $\tilde{\mathbf{x}}_k = \bar{\mathbf{x}}_k$ for unbiased clustering, but it isn't for biased clustering. Hence, we need to appoint mean objects as representatives and recalculate mean objects.

In featureless problem for the set of N members, none of mean objects $\omega(\bar{\mathbf{x}}_k)$ is presented in the distance matrix $D(N, N)$ as a cluster center. We can use the object closest to others in the cluster as the cluster center $\bar{\omega}_k$. In general, if conditions $\tilde{\omega}_k = \bar{\omega}_k$ are true for all clusters, it appears to be the biased clustering, because the center $\mathbf{x}(\bar{\omega}_k)$ isn't the mean object $\bar{\mathbf{x}}_k$ in the unknown feature space.

As it was discussed before [2, 3], featureless versions of k -means algorithm can be represented for distances and similarities and at the same time as k -means and as k -meanless ones.

1. The featureless k -means and k -meanless clustering by distances

The clustering criterion minimizes average of squared distances to cluster centers:

$$J(K) = \frac{1}{N} \sum_{k=1}^K N_k \sigma_k^2 = \sum_{k=1}^K \frac{N_k}{N} \sigma_k^2, \text{ where } \sigma_k^2 = \frac{1}{N_k} \sum_{i=1}^{N_k} (\mathbf{x}_i - \bar{\mathbf{x}}_k)^2 = \frac{1}{N_k} \sum_{i=1}^{N_k} d^2(\mathbf{x}_i, \bar{\mathbf{x}}_k).$$

In the featureless case, the mean objects $\bar{\omega}_k$ provides the unbiased clustering with the cluster dispersion minimized $\sigma_k^2 = \frac{1}{N_k} \sum_{i=1}^{N_k} d^2(\omega_i, \bar{\omega}_k)$, where the criterion $J(K)$ is minimized. If the set Ω will be immersed in a feature space, then two criterions $J^X(K) = \min_{\bar{\mathbf{x}}_1, \dots, \bar{\mathbf{x}}_K} J(K)$ and

$J^D(K) = \min_{\bar{\omega}_1, \dots, \bar{\omega}_K} J(K)$ won't be the same $J^D(K) \geq J^X(K)$ in general. However, $J^X(K) = J^D(K)$, if objects $\mathbf{x}(\bar{\omega}_k)$ and $\bar{\mathbf{x}}_k$ are the same ones. Let's guarantee this condition.

For some $\omega_l \in \Omega$ as a point of the origin and a pair ω_i, ω_j the scalar product is $c_{ij} = (d_{li}^2 + d_{lj}^2 - d_{ij}^2)/2$, where distance is $d_{pq} = d(\omega_p, \omega_q)$ and $c_{ii} = d_{li}^2$ for $i = j$. Therefore, the main diagonal of the matrix $C_l(N, N)$ represents distances squared from the origin $\omega_l \in \Omega$ to other objects. It is convenient [4] to put the origin of the feature space in the "gravity" center of objects $\omega_i \in \Omega$, $i = 1, \dots, N$. According to it, the cluster center $\bar{\omega}_k$ is represented by its distances to other objects without restoring the unknown feature space, where N_k is a number of objects in Ω_k :

$$d^2(\omega_i, \bar{\omega}_k) = \frac{1}{N_k} \sum_{p=1}^{N_k} d_{ip}^2 - \frac{1}{2N_k^2} \sum_{p=1}^{N_k} \sum_{q=1}^{N_k} d_{pq}^2; \omega_p, \omega_q \in \Omega_k,$$

where the cluster dispersion [2, 3] is

$$\sigma_k^2 = \frac{1}{N_k} \sum_{i=1}^{N_k} \left(\frac{1}{N_k} \sum_{p=1}^{N_k} d_{ip}^2 - \frac{1}{2N_k^2} \sum_{p=1}^{N_k} \sum_{q=1}^{N_k} d_{pq}^2 \right) = \frac{1}{2N_k^2} \sum_{p=1}^{N_k} \sum_{q=1}^{N_k} d_{pq}^2. \quad (1)$$

We can represent k -means algorithm in different forms relative the way to recalculate cluster centers in the feature space. Let us develop k -means algorithm for distances only:

a) **Step 0.** Get K representatives $\tilde{\omega}_k^0$, $k = 1, \dots, K$ as may be the most distant to each other objects and put them as centers $\bar{\omega}_k^0$, $k = 1, \dots, K$.

Step s. Reallocate all objects between clusters:

1. Reallocate ω_i to be $\omega_i \in \Omega_k^s$, if $d(\omega_i, \bar{\omega}_k^s) \leq d(\omega_i, \bar{\omega}_j^s)$ for $\omega_i \in \Omega_{j \neq k}^s$, $j = 1, \dots, K$.
2. If needs, recalculate $\bar{\omega}_k^s$, $k = 1, \dots, K$ to represent them by distances $d(\omega_i, \bar{\omega}_k^s)$, $i = 1, \dots, N$.
3. Reallocate next $i = i + 1$ object ω_i .
4. Stop, if it is unbiased clustering with $\tilde{\omega}_k^s = \bar{\omega}_k^s$, $k = 1, \dots, K$, else $\tilde{\omega}_k^{s+1} = \bar{\omega}_k^s$, $s = s + 1$.

According to (1), the average of squared distances between objects in the cluster is the cluster dispersion $\eta_k = \sigma_k^2$. Therefore, the meanless clustering criterion is $\tilde{J}(K) = \frac{1}{N} \sum_{k=1}^K N_k \eta_k$ and $\tilde{J}(K) = J(K)$. If the set Ω will be immersed in a feature space and objects $\mathbf{x}(\bar{\omega}_k)$ and $\bar{\mathbf{x}}_k$ will be the same ones, then two criterions $\tilde{J}^X(K) = \min_{\Omega_1, \dots, \Omega_K \in X} \tilde{J}(K)$ and $\tilde{J}^D(K) = \min_{\Omega_1, \dots, \Omega_K \in D} \tilde{J}(K)$ will be also the same $\tilde{J}^X(K) = \tilde{J}^D(K)$. Let's guarantee this condition for meanless clustering:

b) **Step 0.** Get K subsets Ω_k^0 , $k = 1, \dots, K$ as may be the most compact ones in some way.

Step s. Reallocate all objects between clusters by permutations:

1. Reallocate ω_i to be $\omega_i \in \Omega_k^s$ and $\tilde{J}^s(K) = \tilde{J}_k^s(K)$, if $\tilde{J}_k^s(K) < \tilde{J}_j^s(K)$ for $\omega_i \in \Omega_{j \neq k}^s$, $j = 1, \dots, K$.
2. Reallocate next $i = i + 1$ object ω_i .
3. Stop, if for unbiased clustering none of objects is reallocated, else $s = s + 1$.

2. The featureless k -means and k -meanless clustering by similarities

A positively definite similarity matrix $S(N, N)$ with elements $s_{ij} = s(\omega_i, \omega_j) \geq 0$ can be used as a matrix of scalar products in a metric space of not more than N dimensionality. Relative some point $\omega_k \in \Omega$ as the origin, where $s_{ij} = (d_{ki}^2 + d_{kj}^2 - d_{ij}^2)/2$, $s_{ii} = d_{ki}^2$, distances are defined as $d_{ij}^2 = s_{ii} + s_{jj} - 2s_{ij}$. The cluster center $\bar{\omega}_k$ is represented by its similarities with other objects $\omega_i \in \Omega$, $i = 1, \dots, N$: $s(\omega_i, \bar{\omega}_k) = \frac{1}{N_k} \sum_{p=1}^{N_k} s_{ip}$; $\omega_p \in \Omega_k$. The cluster compactness is the average similarity of the cluster center with other objects in the cluster $\delta_k = \frac{1}{N_k} \sum_{i=1}^{N_k} s(\omega_i, \bar{\omega}_k) = \frac{1}{N_k^2} \sum_{i=1}^{N_k} \sum_{p=1}^{N_k} s_{ip}$; $\omega_i, \omega_p \in \Omega_k$. Unbiased clustering minimizes the cluster dispersion σ_k^2 and maximizes compactness δ_k : $\sigma_k^2 = \frac{1}{2N_k^2} \sum_{i=1}^{N_k} \sum_{j=1}^{N_k} (s_{ii} + s_{jj} - 2s_{ij}) = \frac{1}{N_k} \sum_{i=1}^{N_k} s_{ii} - \delta_k$, and for all clusters:

$$J(K) = \sum_{k=1}^K \frac{N_k}{N} \sigma_k^2 = \sum_{k=1}^K \frac{N_k}{N} \left(\frac{1}{N_k} \sum_{i=1}^{N_k} s_{ii} - \delta_k \right) = \frac{1}{N} \sum_{i=1}^N s_{ii} - \sum_{k=1}^K \frac{N_k}{N} \delta_k = c - \sum_{k=1}^K \frac{N_k}{N} \delta_k. \quad (2)$$

Therefore, for the clustering by similarities we need to maximize weighted compactness $I(K) = \sum_{k=1}^K \frac{N_k}{N} \delta_k$, where $I(K) = c - J(K)$. We can develop k -means algorithm by similarities to build unbiased clustering relative centers:

c) **Step 0.** Get K representatives $\tilde{\omega}_k^0$, $k = 1, \dots, K$ as may be the least similar to each other objects and put them as centers $\bar{\omega}_k^0$, $k = 1, \dots, K$.

Step s. Reallocate all objects between clusters:

1. Reallocate ω_i to be $\omega_i \in \Omega_k^s$, if $s(\omega_i, \bar{\omega}_k^s) \geq s(\omega_i, \bar{\omega}_j^s)$ for $\omega_i \in \Omega_{j \neq k}^s$, $j = 1, \dots, K$.
2. If needs, recalculate $\bar{\omega}_k^s$, $k = 1, \dots, K$ and represent them by similarities $s(\omega_i, \bar{\omega}_k^s)$, $i = 1, \dots, N$.
3. Reallocate next $i = i + 1$ object ω_i .
4. Stop, if it is unbiased clustering with $\tilde{\omega}_k^s = \bar{\omega}_k^s$, $k = 1, \dots, K$, else $\tilde{\omega}_k^{s+1} = \bar{\omega}_k^s$, $s = s + 1$.

In addition, we can develop k -meanless algorithm to build unbiased clustering by similarities:

d) **Step 0.** Get K subsets Ω_k^0 , $k = 1, \dots, K$ as may be the most compact ones in some way.

Step s. Reallocate all objects between clusters by permutations:

1. Reallocate ω_i to be $\omega_i \in \Omega_k^s$ and $I^s(K) = I_k^s(K)$, if $I_k^s(K) > I_j^s(K)$ for $\omega_i \in \Omega_{j \neq k}^s$, $j = 1, \dots, K$.
2. Reallocate next $i = i + 1$ object ω_i .
3. Stop, if for unbiased clustering none of objects is reallocated, else $s = s + 1$.

3. The featureless k -meanless clustering with the two-part objective function

We develop here the new version of k -means algorithm again based on the bi-partial approach to clustering [5, 6]. In contrast to the idea of combining different scales to represent intercluster similarity with intracluster distances or vice versa [5, 6], we don't need to find the linear combination coefficient, because distances and similarities are presented in the same unknown metric space.

For meaningless clustering according to (1) with dispersions η_k we need to minimize the criterion $\tilde{J}(K) = \frac{1}{2N} \sum_{k=1}^K \frac{1}{N_k} \sum_{p=1}^{N_k} \sum_{q=1}^{N_k} d_{pq}^2$. In addition, for the dual problem according to (2) we need to maximize the criterion $I(K) = \frac{1}{N} \sum_{k=1}^K \frac{1}{N_k} \sum_{p=1}^{N_k} \sum_{q=1}^{N_k} s_{pq}$.

In the first case according to the bi-partial approach, a two-part objective function $\tilde{J}_\delta(K) = \tilde{J}(K) + \delta(K)$ combines intracluster distances $\tilde{J}(K)$ with intercluster similarity $\delta(K)$ and needs to be minimized. Let's define the intercluster similarity $\delta(K) = \frac{1}{K} \sum_{k=1}^K s(\bar{\omega}_k, \bar{\omega}_0)$ relative the center of the whole set as the object $\bar{\omega}_0$, represented by its similarities $s(\bar{\omega}_k, \bar{\omega}_0) = \frac{1}{K} \sum_{p=1}^K s(\bar{\omega}_k, \bar{\omega}_p)$ with other centers $\bar{\omega}_k, k=1, \dots, K$. Therefore, $\delta(K) = \frac{1}{K^2} \sum_{k=1}^K \sum_{l=1}^K s(\bar{\omega}_k, \bar{\omega}_l)$.

Let's define the similarity $s(\bar{\omega}_k, \bar{\omega}_l)$ between cluster centers. Each cluster center $\bar{\omega}_k, k=1, \dots, K$ is represented by similarities with other objects $\omega_i, i=1, \dots, N$ as $s(\omega_i, \bar{\omega}_k) = \frac{1}{N_k} \sum_{p=1}^{N_k} s_{ip}$, $\omega_p \in \Omega_k$. According to it, the similarity between cluster centers is $s(\bar{\omega}_k, \bar{\omega}_l) = \frac{1}{N} \sum_{i=1}^N s(\omega_i, \bar{\omega}_k) s(\omega_i, \bar{\omega}_l) = \frac{1}{NN_k N_l} \sum_{i=1}^N \sum_{p=1}^{N_k} s_{ip} \sum_{q=1}^{N_l} s_{iq}$. As a result, the intercluster similarity is:

$$\delta(K) = \frac{1}{NK^2} \sum_{k=1}^K \sum_{l=1}^K \frac{1}{N_k N_l} \sum_{i=1}^N \sum_{p=1}^{N_k} s_{ip} \sum_{q=1}^{N_l} s_{iq}, \text{ where } \omega_p \in \Omega_k, \omega_q \in \Omega_l.$$

In addition, we can see, that the cluster center $\bar{\omega}_k$ is represented by its similarities with other objects $\omega_i \in \Omega$ and, specifically, with objects from other cluster $\omega_i \in \Omega_l$ only. Therefore, we can get the average similarity of objects $\omega_i \in \Omega_l$ with the cluster center $\bar{\omega}_k$: $s(\Omega_l, \bar{\omega}_k) = \frac{1}{N_l} \sum_{i=1}^{N_l} s(\omega_i, \bar{\omega}_k) = \frac{1}{N_l} \sum_{i=1}^{N_l} \frac{1}{N_k} \sum_{p=1}^{N_k} s_{ip}$, $\omega_p \in \Omega_k$. It is evident, that $s(\Omega_l, \bar{\omega}_k) = s(\Omega_k, \bar{\omega}_l)$. Therefore, we can use the notation $s(\Omega_l, \bar{\omega}_k) = s(\Omega_k, \bar{\omega}_l) = s(\Omega_l, \Omega_k)$ to denote another type of pairwise intercluster similarity. As a result, another type of the intercluster similarity is:

$$\delta(K) = \frac{1}{K^2} \sum_{k=1}^K \sum_{l=1}^K \frac{1}{N_k N_l} \sum_{p=1}^{N_k} \sum_{q=1}^{N_l} s_{pq}, \text{ where } \omega_p \in \Omega_k, \omega_q \in \Omega_l.$$

As a result, we can use the version b) of k -meanless clustering for the two-part objective function $\tilde{J}_\delta(K) = \tilde{J}(K) + \delta(K)$ instead of the criterion $\tilde{J}(K)$. In addition, for the dual problem we can use the corresponding two-part objective function $I_\sigma(K) = I(K) + \sigma^2(K)$ with intercluster dispersion $\sigma^2(K)$ for the version d) of k -meanless clustering instead of the criterion $I(K)$.

Conclusion

It is evident, that we need to reduce the computational complexity of proposed algorithms based on increment of the criterion function. We can do it using optimal techniques.

This work is supported by the RFBR Grant 13-07-00010. I would like to thank Jan W. Owsinski for helpful discussion of the bi-partial approach.

References

1. Featureless pattern recognition in an imaginary Hilbert space and its application to protein fold classification / S.D. Dvoenko [et al.] // 2nd Intern. Workshop on Machine Learning and Data Mining in Pattern Recognition (MLDM-2001). – Leipzig, 2001. – P. 322–336.
2. Dvoenko, S.D. Clustering of a set of objects / S.D. Dvoenko // Proc. of the 9th Intern. Conference on Pattern Recognition and Information Processing (PRIP-2007). – Minsk, 2007. – Vol. 1. – P. 93–97.

3. Dvoenko, S.D. On clustering of a set of members by distances and similarities / S.D. Dvoenko // Proc. of the 11th Intern. Conference on Pattern Recognition and Information Processing (PRIP-2011). – Minsk, BSUIR, 2011. – P. 104–107.
4. Torgerson, W.S. Theory and methods of scaling / W.S. Torgerson. – Wiley, N.Y., 1958.
5. Owsinski, J.W. The bi-partial approach in clustering and ordering: the model and the algorithms / J.W. Owsinski // Statistica & Applicazioni. – 2011. – Special Issue. – P. 43–59.
6. Owsinski, J.W. The matter of scale: perceiving distances and proximities in the bi-partial clustering setting / J.W. Owsinski // Abst. of the Conference of the International Federation of Classification Societies (IFCS-2013). – Tilburg, the Netherlands, 2013. – P. 88–89.

CREATING OF THE DISCRETE ORTHOGONAL TRANSFORMATIONS WITH THE USE OF CELLULAR AUTOMATA DYNAMICS

O. Evsutin, A. Shelupanov
Tomsk State University of Control Systems and Radioelectronics, Russia
e-mail: eoo@keva.tusur.ru

The dynamics of one-dimensional cellular automaton is applied for the construction of discrete orthogonal transformation bases which can be used for generation of digital image compression algorithms. We suggest a comparison technique of such transformations in respect of noises shown as a result of information losses on the restored data elements.

Introduction

Currently, various applications of cellular automaton theory to digital image processing tasks are known. The given mathematical apparatus is used for images enhancement [1, 2], segmentation [3, 4], edge extraction and text recognition [5, 6], construction of secret sharing schemes, based on the use of digital images [7], and also in computer steganography for the embedding of water marks into digital images [8].

In [9] digital image compression method is considered, based on discrete orthogonal transformations, which bases are built from the evolving states of classical cellular automaton with Moore and von Neumann neighborhood. However, such transformations are a particular case of a famous Walsh-Hadamard transformation due to the use of internal states binary alphabet.

The given work describes generalization of the approach of discrete orthogonal transformation creating with the use of cellular automata dynamics.

1. Cellular automaton with code set

Let us describe a mathematical model of cellular automaton as a set of components $CA = \langle Z^n, \mathbf{L}, A, \mathbf{Y}, \sigma \rangle$, where Z^n – is a space of lattice cells coordinates; $\mathbf{L} = (l_1, K, l_n)$, $l_i > 0$, $i = \overline{1, n}$ – vector that sets the lattice size; A – internal state alphabet, that determines the finite set of values of a single cell; \mathbf{Y} – cell neighborhood, represented as a relative indexes vector, which determines similar for each lattice cell number and position order of neighboring cells, i.e. cells, the current states of which will influence the state of a given cell in the next point of time; σ – local function of transition, defined analytically or given by the set of parallel substitutions, and applied to all lattice cells simultaneously. The set of arguments of the given function is defined by \mathbf{Y} neighborhood.

As extension of classical cellular automaton model we will introduce a notion of cellular automaton with code set $CA_K = \langle CA, K, \varphi \rangle$, where CA is a cellular automaton with the internal states alphabet A , K – an ordered set of values, so that $|K| = |A|$, and function $\varphi: A \rightarrow K$ assigns code set elements K to symbols of internal states alphabet A . An introduced extension of a classical model is used for formalization of the approach aimed at generation of code sequence of the given type from the evolving states of cellular automaton.

2. Creating of orthogonal bases with the use of code set cellular automaton

For the creation of orthogonal basis we should take a one-dimensional cellular automaton CA with lattice length N , define a code set cellular automaton $CA_K = \langle CA, K, \varphi \rangle$ over it and for some initial lattice state study the evolution history of cellular automaton – a lattice states sequence at time points $t = 1, 2, K$. The first basis vector generates from the initial lattice state by means of function $\varphi: A \rightarrow K$, afterwards more pairwise orthogonal $N-1$ vectors are chosen from the cellular automaton evolution history. The respective algorithms are considered in [10].

The set of orthogonal bases generated from the evolving states of code set cellular automaton $CA_K = \langle CA, K, \varphi \rangle$ for all possible initial lattice states is referred to as bases family and denoted by $\Sigma(CA_K)$.

It has been found experimentally that certain basis families, in general case, can contain hundreds of thousands of different bases (experiments were carried out for $N = 8$), but only some of them are of practical interest. Therefore the selection task of the best bases among the great number of generated ones has appeared.

Let us carry out the investigation of a single basis family as follows. First of all from the given family it is necessary to single out subfamily of bases that define transformations, have similar characteristics of the interesting to us type.

For this purpose, let us set the vector of values with spatial redundancy and transform it using the formula $\mathbf{G} = \mathbf{F}(\mathbf{D} \cdot \mathbf{C})$, where $\mathbf{F} = (f_i)_{i=1}^N$ — set vector; $\mathbf{C} = (c_{ij})_{i=1, j=1}^{N, N}$, $c_{ij} \in K$, $i, j = \overline{1, N}$ — discrete orthogonal transformation basis; $\mathbf{D} = (d_{ij})_{i=1, j=1}^{N, N}$ — diagonal normalization matrix.

After vector \mathbf{F} transformation it is necessary to define, which of the transformed values may be related to low-frequency components containing the main (average) information about the initial vector \mathbf{F} , and which may be related to high-frequency components indicating the differences among separate elements of vector \mathbf{F} . For that purpose we will introduce λ parameter, defining minimal relation of element g_i , $i = \overline{1, N}$, value of vector \mathbf{G} to $M = \frac{1}{N} \sum_{j=1}^N f_j$ value, at which the given element is considered as low-frequency component. Then lets calculate the variation value of low-frequency components values using the formula $\nu = \frac{\min(\hat{g}_i)}{\max(\hat{g}_i)}$, where \hat{g}_i , $i = \overline{1, r}$, $0 \leq r \leq N$ — are elements of vector \mathbf{G} , that are considered as low-frequency components.

The respective algorithm is described in [11].

The set of orthogonal bases selected from the family $\Sigma(\text{CA}_K)$ at given values of low-frequency components number r , coefficient λ , and variation parameter ν , will be referred to as a subfamily and denoted by $\Sigma_{r, \lambda, \nu}(\text{CA}_K)$, $\Sigma_{r, \lambda, \nu}(\text{CA}_K) \subseteq \Sigma(\text{CA}_K)$.

In this paper, orthogonal transformations similar to discrete wavelet transformation, where $N \equiv 0 \pmod{2}$ and $r = \frac{N}{2}$ ($N = 8$), were built and considered. These transformations were built with the use of cellular automaton $\text{CA}_K = \langle \text{CA}_p, K_s, \varphi \rangle$ with code set $K_s = \{-s, -s+1, s-1, s\}$, defined over block cellular automaton [12] with internal states alphabet $A = \{0, 1, 2, 3\}$ and bijective block function of transition.

3. Study of effectiveness of discrete orthogonal transformations

Let us then consider the possibility of application of the obtained orthogonal transformations for the digital image compression at the stage of pixel value decorrelation. As an orthogonal transformation efficiency criterion we will take the number of zero values among the transformed data elements after their quantization in conditions that the initial data had spatial redundancy. In order to calculate the information losses produced during this process we will use the root-mean-square error

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (f_i - g_i)^2}.$$

For the selection of the best orthogonal transformations according to the restoration error we suggest the following technique.

1. Generation of the set (subfamily or set of subfamilies) Σ of orthogonal bases of N order with the use of code set cellular automaton dynamics.

2. Generation of the set of integer matrixes (simple test images) describing some basic behavior features of certain parts of a digital image: gradual color blend, presence of some differentiating areas

with gradual color blend, presence of the object on the background of gradual color blend, presence of lines on the background of gradual color blend.

3. Application of orthogonal transformation on the base of each of the built bases $\mathbf{C} \in \mathcal{Z}$ to each of the given test images with the following quantization of transformed data elements using a simple scheme with two coefficients $SQ = (q_L, q_H)$, where q_L – is a low-frequency components quantizer, q_H – a high-frequency components quantizer.

4. Inversion of each transformation with calculation of RMSE value and of zeros number among the transformed data elements in each image and calculation of the following statistical characteristics: average number of zeros among the quantized data elements, variation coefficient of the given value, average RMSE value and variation coefficient of the given value, which for some basis $\mathbf{C} \in \mathcal{Z}$ we will denote respectively by $z(\mathbf{C})$, $V_z(\mathbf{C})$, $RMSE(\mathbf{C})$, $V_{RMSE}(\mathbf{C})$.

5. Partition of the \mathcal{Z} set into l disjoint subsets $\mathcal{Z} = \bigcup_i \mathcal{Z}_i$, where $\mathcal{Z}_i, i = \overline{1, l}$, is defined in the following way: the number line segment $\left[0, \frac{3}{4} N^2\right]$, composing a set of possible characteristic values z , is divided into l parts, and if $z(\mathbf{C}), \mathbf{C} \in \mathcal{Z}$, belongs to segment with i index, then $\mathbf{C} \in \mathcal{Z}_i$.

6. Selection of the best orthogonal bases from each subset $\mathcal{Z}_i, i = \overline{1, l}$, the average RMSE value of which is minimal, and their integration into the set $\hat{\mathcal{Z}} \subset \mathcal{Z}$. When considering such bases $\mathbf{C}_1, \mathbf{C}_2 \in \mathcal{Z}$, for which $\mathbf{C}_1, \mathbf{C}_2 \in \mathcal{Z}$, a best basis is that one with a smaller RMSE value variation.

7. Application of orthogonal transformations with bases from $\hat{\mathcal{Z}}$ to each of the given simple test images so, that the information loss level was similar (this is expressed by approximately equal number of zeros among the transformed data elements).

8. Final selection of the best orthogonal bases according to the restoration error.

Fig. 1, 2 shows typical relation between $z(\mathbf{C})$ and $RMSE(\mathbf{C})$ characteristics for bases of $\Sigma_{4, 0.5, 0.8}(CA_{K_s})$ subfamilies for $s = \overline{2, 200}$: with the increase of the number of zeros among the transformed data elements the reduction of restoration error is observed.

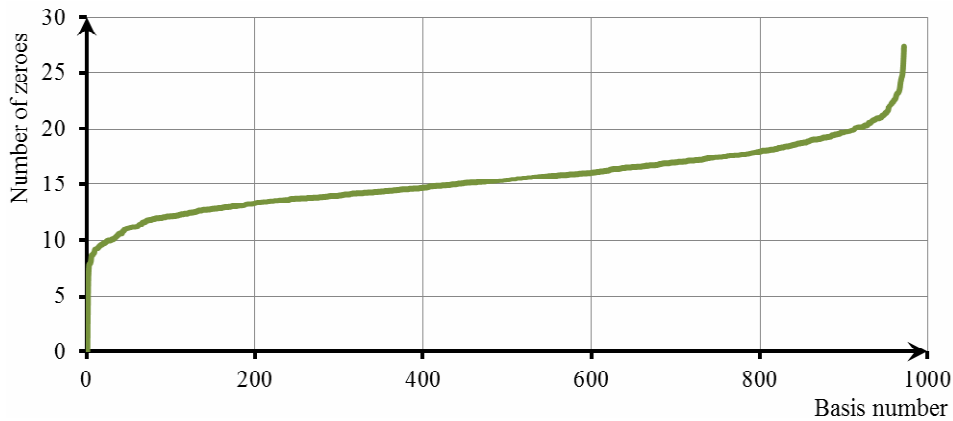


Fig. 1. Characteristics $z(\mathbf{C})$ of subfamily bases for $s = 4$

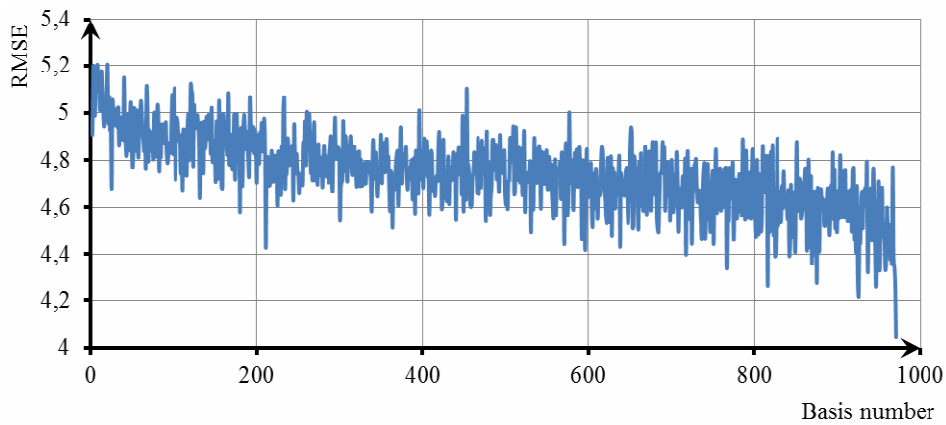


Fig. 2. Characteristics RMSE(C) of subfamily bases for $s = 4$

4. Example

As an example let us consider basis $\mathbf{C} =$

$$\begin{bmatrix} 3 & -3 & -4 & -4 & 3 & 3 & 4 & 4 \\ 3 & 4 & 3 & 4 & -4 & -3 & 4 & 3 \\ 4 & -3 & -4 & 3 & 3 & -4 & 3 & -4 \\ 4 & -4 & 3 & -3 & -4 & 4 & 3 & -3 \\ -3 & 3 & 4 & -4 & 3 & -3 & 4 & -4 \\ -3 & -4 & 3 & 4 & 4 & 3 & 4 & 3 \\ 4 & 4 & 3 & 3 & 4 & 4 & -3 & -3 \\ 4 & -3 & 4 & -3 & 3 & -4 & -3 & 4 \end{bmatrix}$$
 from

$\Sigma_{4,0.5,0.8}(\text{CA}_{K_4})$ subfamily, obtained with the dynamics of code set cellular automaton $K_4 = \{-4, -3, 3, 4\}$. The given basis is the best representative of its subfamily according to restoration error. Fig. 3, a shows the result of the respective transformation of the classic test image Lenna with total removal of all 1-level high-frequency components (that is 75% of all transformed data elements). This compares to the shown in fig. 3, b result of the similar Lenna image transformation on the base of random representative of the same subfamily. It can be seen, that image (fig. 3, a) doesn't have any visible artifacts, while the image (fig. 3, b) showed typical grid.



Fig. 3. The result of Lenna image transformation: using basis with the best characteristics (a), using random basis (b)

Conclusion

In conclusion we would like to point out, that the given approach to the designing of discrete orthogonal transformations with cellular automaton dynamics in general case assumes the use of code sets of any cardinality. However, in practice the use of code sets, consisting of more than six elements, is difficult at present time because of the considerable calculating time. Therefore, our subsequent work will be focused on two main areas: working with code sets type differing from the type considered in the given work, and the use of cellular automata of higher order (the order of cellular automaton can be defined by its internal states alphabet cardinality or by binary logarithm of the given value). Moreover, it should be noted that the obtained results can be applied to solving of information security tasks, particularly those, where transformation of data sets of different dimension are required [13–15].

References

1. Rosin, P.L. Training cellular automata for image processing / P.L. Rosin // Proc. of the 14th Scandinavian conf. on Image Analysis (SCIA 2005). – Joensuu, Finland, 2005. – P. 195–204.
2. Panagiotopoulos, F.K. Quantum-dot cellular automata design for median filtering and mathematical morphology operations on binary images / F.K. Panagiotopoulos, V.A. Mardiris, V. Chatzis // Proc. of the 10th Intern. Conf. on Cellular Automata for Research and Industry (ACRI 2012). – Santorini Island, Greece, 2012. – P. 554–564.
3. Kauffmann, C. Seeded ND medical image segmentation by cellular automaton on GPU / C. Kauffmann, N. Piché // Intern. Journal of Computer Assisted Radiology and Surgery. – 2010. – Vol. 5 (3). – P. 251–262.
4. An effective approach of lesion segmentation within the breast ultrasound image based on the cellular automata principle / Y. Liu [et al.] // Journal of Digital Imaging. – 2012. – Vol. 25 (5). – P. 580–590.
5. Chen, Y. A cellular automatic method for the edge detection of images / Y. Chen, Z. Yan // Proc. of the 4th Intern. Conf. on Intelligent Computing (ICIC 2008). – Shanghai, China, 2008. – P. 935–942.
6. Text extraction and enhancement of binary images using cellular automata / G. Sahoo [et al.] // Intern. Journal of Automation and Computing. – 2009. – Vol. 6 (3). – P. 254–260.
7. del Rey, A.M. A secret sharing scheme for digital images based on two-dimensional linear cellular automata / A.M. del Rey // Proc. of the 12th Intern. workshop Combinatorial Image Analysis (IWCIA 2008). – Buffalo, NY, USA, 2008. – P. 318–329.
8. A new JPEG image watermarking algorithm based on cellular automata / H. Wu [et al.] // Journal of Information & Computational Science. – 2011. – Vol. 8 (12). – P. 2431–2439.
9. Lafe, O. Data compression and encryption using cellular automata transforms / O. Lafe // Engineering Applications of Artificial Intelligence. – 1997. – Vol. 10 (6). – P. 581–591.
10. Evsutin, O.O. Designing and testing of the computational approach of decorrelating transformation bases construction with the use of partitioning cellular automata / O.O. Evsutin, S.K. Rososhek // Trudy SPIIRAN. – 2012. – Vol. 4 (23). – P. 324–342. [In Russian]
11. Evsutin, O.O. Applications of cellular automata in the field of information security and data processing / O.O. Evsutin, A.A. Shelupanov // Doklady TUSUR. – 2012. – Vol. 1–2 (25). – P. 119–125. [In Russian]
12. Toffoli, T. Cellular automata machines / T. Toffoli, N. Margolus. – M. : Mir, 1991. – 280 p. [In Russian]
13. Cryptographic protocol in limited resource systems / R.V. Meshcheryakov [et al.] // Computational Technologies. – 2007. – Vol. 12 (S1). – P. 51–61. [In Russian]
14. Mironova, V.G. Intruder model the security of confidential information / V.G. Mironova, A.A. Shelupanov // Information Science and Control Systems. – 2012. – Vol. 1 (31). – P. 28–35. [In Russian]
15. Prishchep, S.V. Approaches and criteria of information security risk assessment / S.V. Prishchep, S.V. Timchenko, A.A. Shelupanov // Bezopasnost' informacionnyh tehnologij. – 2007. – Vol. 4. – P. 15–21. [In Russian]

RECOGNITION OF ZERO BITS OF 3-SAT PROBLEM BY APPLYING LINEAR ALGEBRA'S METHODS¹

R.T. Faizullin, Y.Y. Ogorodnikov
Omsk State Technical University, Russia
e-mail: r.t.faizullin@mail.ru, yogorodnikov@gmail.com

The paper represents two methods of recognizing zero bits of 3-SAT problem. The first is based on the reduction of the satisfiability problem to an equivalent problem of minimizing a continuous smooth function by method of successive approximations, extended by using permutation's matrix. Another way is to reduce the system of linear algebraic equations with symmetric diagonally dominant matrix.

Introduction

The problem of satisfiability of Boolean formulas (SAT) occupies a central place in the complexity theory, an important branch of mathematics. In 1971 S. Cook proved, that SAT problem in CNF-normal form is NP-complete, i.e. all other problems of NP class are reduced to SAT in polynomial time. In particular, factorization problem has a special meaning for cryptography [1]. The reason for this interest is that there is no effective algorithm of solving this problem. Nowadays there is well-known reduction of this problem to equivalent SAT instances [2]. Thus, the part of performing set for 3-SAT will match the decomposition of number for equivalent factorization problem. Obviously, if bits of this part will be recognized with high probability (≥ 0.9), it can be possible to attack RSA algorithm. The instance of SAT problem of small dimensions can be solved by backtracking methods, but with increasing the dimension the time of solution grows exponentially. This is the reason to research more effective algorithms. The complex study is very often used, because the common results by various approaches may sufficiently increase the probability of bit's recognizing. This article deals with modification of one hybrid algorithm already published [3] by using the methods of linear algebra. Also, a new way based on the concepts of linear algebra is proposed.

1. Method of successive approximations with permutations

In papers [3, 5] the method of reducing 3-SAT problem to equivalent problem of minimization continuous smooth function is presented. The result system has the following form:

$$\left(\sum_{k \in T(j)} \prod_{l=1, l \neq j}^N (p_{k,l}(x)) x_j - \sum_{k \in Q(j)} \prod_{l=1, l \neq j}^N p_{k,l}(x) = 0 \quad (j = 1, \dots, N), \right. \quad (1)$$

where $T(j) = \{k \in \{1, \dots, M\} : p_{k,j}(x) = x_j^2 \text{ or } (1-x_j)^2\}$ and $Q(j) = \{k \in \{1, \dots, M\} : p_{k,j}(x) = (1-x_j)^2\}$.

Let's convert the resulting system of equations (1) to form:

$$A^* x = B^*, \quad (2)$$

where A^* – the coefficient matrix; B^* – column of free terms. With the method of successive approximations applying to (2), the vector approximation $x^* = (x_1, x_2, \dots, x_n)$ is obtained. After some tests were made, it was found that some bits of approximation with sufficiently high probability (greater than or equal to 0.9) coincide with the bits of the exact solution. Let's investigate the modification of this method to increase the number of such bits. For this purpose the methods of linear algebra should be used.

¹ Supported by RFBR, research project No. 2-07-00294-a

Take a permutation $\sigma = \begin{pmatrix} 1 & 2 & \dots & n \\ \sigma(1) & \sigma(2) & \dots & \sigma(n) \end{pmatrix}$ order n , where n is the dimension of A^* .

Denote $P_\sigma = \begin{pmatrix} e_{\sigma(1)} \\ e_{\sigma(2)} \\ \mathbf{M} \\ e_{\sigma(n)} \end{pmatrix}$ as corresponding permutation's matrix, where $e_{\sigma(i)}$ - vector with dimension n , i -th element of which equals 1, the other ones equals zero [1].

Apply the matrix P_σ to matrixes A^* and B^* by multiplying it on the left side. After the multiplying the system (2) takes a form:

$$P_\sigma A^* X = P_\sigma B^*; \quad (3)$$

$$A_\sigma^* X = B_\sigma^*. \quad (4)$$

With the method of successive approximations applying to the system (4), the vector $x_\sigma^* = (x_1, x_2, \dots, x_n)$ is obtained. It's not difficult to see that generally vectors x^* and x_σ^* are various. Use permutation σ and corresponding matrix P_σ for accumulation of statistics used in the calculation of the probability of the coincidence of approximation bits with the bits of exact solution (further we will named such probability a frequency stability). To do this, apply a permutation matrix to the system (4), then the method of successive approximations should be used. It should be repeated M times for generating statistics, where M is the number of instances of SAT-problem. After this procedure the frequencies vector $\nu_{\sigma(i)} = (\nu_1, \nu_2, \dots, \nu_n)$ would be formed. Notice, that if we do the same steps for various permutations, we obtain the set of frequencies vector $V = \{\nu_i\}$. These data can be used for the modification of the method of successive approximations.

Algorithm 1. Searching for approximate solution of system equations that equivalent 3-SAT-problem, using a set of frequencies vector stability bit.

Input data: A set of frequencies vector $V = \{\nu_i\}$.

Step 1. Make the transition from satisfiability problem to minimizing continuous function.

Step 2. Apply to the system equations from previous step the method of successive approximations.

Marked result as $x = (x_1, x_2, \dots, x_n)$.

Step 3. Determine functions $Q^0(\nu_i, x_i) = \begin{cases} \nu_i, & x_i = 0 \\ 1 - \nu_i, & x_i = 1 \end{cases}$ and $Q^1(\nu_i, x_i) = \begin{cases} \nu_i, & x_i = 1 \\ 1 - \nu_i, & x_i = 0 \end{cases}$

Step 4. Calculate sums $y_j^0 = \frac{1}{N_\pi} \sum_{i=1}^{N_\pi} Q^0(\nu_{i_j}, x_j)$ and $y_j^1 = \frac{1}{N_\pi} \sum_{i=1}^{N_\pi} Q^1(\nu_{i_j}, x_j)$, where $j = 1..n$.

Step 5. Form components of new approximations $\tilde{x}_j = \begin{cases} 1, & y_j^1 \geq y_j^0 \\ 0, & \text{otherwise} \end{cases}$, $j = 1..n$.

Output data: Vector-approximation $\tilde{x} = (\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_n)$.

End of Algorithm 1.

There is an alternative way to determine the probability of equality bit zero, based on reducing the problem to the SAT system of linear equations [4].

2. Diagonal method for determining a frequency stability

Let's get CNF:

$$L(x) = \prod_{i=1}^M C_i, \quad (5)$$

where C_i is disjunction like $\vee q_{i,j}$. Here $q_{i,j} = x_j$ or $q_{i,j} = \overline{x_j}$.

Transition from Boolean variables to real is performed according to formulas: $x_j \rightarrow y_j$, $\overline{x_j} \rightarrow 1 - y_j$. Thus there is a transition to a system of linear equations

$$Ay = f. \quad (6)$$

Obviously, the right-hand side of the i -th equation f_i equals to the number $q_{i,j}$ taking the values "TRUE" in the original CNF. Suppose that the vector f is known for us, for example $\forall j, f_j = 2$, and consider the question: is there a solution for the system (6), components of which are equal 0 or 1? Transform the system according to the formula:

$$RAy = By = Rf = g, \quad (7)$$

where R is generally an arbitrary rectangular matrix of size $N \times M$ (N is the total number of variables, M is the total number of clauses). But the direct selection of the matrix R and its multiplication by the matrix A right does not always yield a matrix that is convenient for further research and operations. One of convenient type of matrix is sparse matrix. Therefore, instead of choosing the matrix R we will build a sparse matrix B of dimension N , while taking care not to lose the certainty inherent in the original problem.

Consider a variable index j and the corresponding equation (6). We will add these equations, accumulating unknown in the j -th row of the matrix B , multiplying by -1 if the variable enters the equation with a minus sign. It does not guarantee us the certainty of the original matrix, but lets not lose sparsity structure inherent in the matrix A . It is easy to show that the matrix B is symmetric.

In general, the prevalence is not strict, however, in practice more than the diagonal elements of the sum of the elements in the corresponding row. Now we turn our attention to the right side g . It should be build synchronously with matrix B . Obviously, it is obtained by adding a components of column vector f . Above it has been shown that $f_j = 1 \vee 2 \vee 3$. Denote as g^2 the vector of right part, resulting the conversion f to g when $\forall i f_i = 2$. The vector $g = Rf$, unknown in general case, can be presented in form $g^2 + \Delta g$, i.e. as sum of a vector corresponds constants with vector "perturbation". Try to estimate the norm of the vector of perturbations Δg about deposits in g^2 from $f_j \neq 2$. Components of vector Δg obtained by summing all disjunctions containing literal j , and addition can arrange dichotomous, binary tree climbing to the top amount. Consider all cases summation at the lowest level when at least one of the logical equations worth unsigned literal denial. When pairwise accumulation equations corresponding to the logical, right sides can be folded only one of the following ways:

$$(2+1)+(2+1), (2+1)+(2-1), (2+1)-(2+1), (2+1)-(2-1), 2+(2-1), 2+(2+1), 2-(2-1), 2-(2+1).$$

We assume these cases are equally probable, for example, this assumption is justified for randomly generated CNF, and try to determine how changes in the average component g_j compared with the component g_j^2 . Obviously, the expectation would be perturbations power $1/4$, while the right part undisturbed value equals 2. We can conclude that the "perturbation" really can be treated as a perturbation on the average comparable and less the norm than the original vector g^2 .

Consider the maximal components of solutions \overline{y} of the system (7) with right side g^2 , i.e. those components which are substantially greater than unity. It is understood that Δg corrects an unknown disturbance vector \overline{y} to a real component of the vector solutions y having whole components 0 and 1. Due to the relative smallness of the perturbation of the right over the original right-hand side is more likely an adjustment to the maximum component units, and not up to scratch. In favor of such an adjustment is the fact that the matrix - matrix diagonally dominant.

3. Numerical experiments

This section presents the results of testing the method of successive approximations with permutations and diagonal method. Experiments were performed on 3-CNF equivalent to factorization problem. It is ensured that all 3-CNF are different and have unique satisfying assignment. Table 1 shows the results of testing the method of successive approximations with permutations for 3-SAT problem, equivalent to factoring the number of dimension 100, 200, 300, 400 bits, respectively.

Table 1

Test results for method of successive approximations with permutations

Dimension of the problem, bit	Number of variables of equivalent 3-SAT	The number of zero bits determined by algorithm 1 N_1	The number of true zero bits N_2	Ratio N_1 to N_2
100	14700	598	646	0.925
200	59400	2677	3075	0.8705
300	134100	6215	7042	0.8825
400	238800	10961	12442	0.881

There were 100 random permutations applied for each dimension factorisable number. In turn, for each permutation used 400 random formulations of the problem SAT, produced using high-quality random number generator. As seen from the results, the attitude to remain constant with increasing dimension of the problem. Table 2 shows the results of testing the diagonal method for CNF equivalent to the factorization problem. Here, the dimension of the problem means the size of factorable number $n = pq$.

Table 2

Test results for diagonal method

Dimension of the problem, bit	Number of variables of equivalent 3-SAT	The number of zero bits determined by diagonal method N_1	The number of true zero bits N_2	Ratio N_1 to N_2
100	14700	4277	4947	0.8645
200	59400	17136	19894	0.8613
300	134100	38942	44828	0.8686
400	238800	68849	79786	0.8629

This method defines more number of true zero bits. Strong point of this method is that the ratio N_2 to N_1 remains constant with increasing dimension of the problem.

Obviously, by combining above-considered methods may obtain a sufficiently large number of bits with high frequency stability. Zero bits, determined by these methods, can be used in the preparation of a new, closer to the exact hamming approach and to reduce (eliminate) CNF by deleting clauses containing literals whose values are determined with high probability.

Conclusion

This article describes two approaches to the search for a satisfying assignment problem SAT, by using the methods of linear algebra. The first of these is to expand the known method of minimizing the functional by adding permutation matrices. Another method is based on creating a matrix diagonally dominant structure using CNF. The numerical experiments are made, the various variants of application are proposed.

In the future it plans to develop an algorithm for determining the factors bit factorization problem, which has the starting approximation containing bits which values are determined with high probability, will find the bit values of one of the factors.

References

1. Costrickin, A.I. Introduction to linear algebra. Basis of algebra / A.I. Costrickin. – M. : Fismatlit, 1994. – P. 59–71.
2. Dulkey, V.I. CNF for factorization, discrete logarithm and discrete logarithm on elliptic curve problems / V.I. Dulkey // Dissertation for candidate of physical-mathematics sciences. – Omsk, 2010. – P. 9–22.

3. Dulkey, V.I. A hybrid method of searching the approximate solution of 3-SAT, associated with factorization problem / V.I. Dulkey, R.T. Faizullin, Y.Y. Ogorodnikov // Proceedings of the Institute of Mathematics and Mechanics URAN. – Ekaterinburg, 2013. – Vol. 19 (2). – P. 285–294.
4. Faizullin, R.T. Problems of linear algebra, concerning the SAT problem / R.T. Faizullin // PDM. – 2009. – Appl. 1. – P. 90–91.
5. Hnikin, I.G. The minimization of functionals associated with satisfiability problems / I.G. Hnikin // Dissertation for candidate of technical sciences. – Omsk, 2009. – P. 38–42.

DECISION SUPPORT FOR PRECISION FARMING COMPLEX

V. Ganchenko, A. Doudkin, A. Petrovsky
United Institute of Informatics Problems of the NAS of Belarus, Minsk
e-mail: ganchenko@lsi.bas-net.by

In the article algorithms for decision support for hardware and software complex are described. The complex is used for few precision farming tasks: data mining, data processing, decision making and control of fertilizers applying. The complex is designed to reduce costs and environmental burden on potato. The complex is based on processing aerial images photographs of potato fields.

Introduction

Modern methods of thematic cartography and resource management are demanded more and more for remote sensing data (RSD) processing. These data are used in areas such as cartography and land cadastre [1], agronomy and precision agriculture [2, 3], a forestry [4], a development of water systems [5], an environmental monitoring [6] etc.

One of the most important areas of image processing is a precision agriculture area. Efficient processing of raw data allows reducing material and other costs in problems associated with crop cultivation and forecasting, a monitoring of level of crops germination and many other applications.

The basic concept of precision agriculture is the fact that a vegetation cover is not uniform within a single field. Up-to-date technologies are used to evaluate and detect these irregularities: global positioning systems (GPS, GLONASS), special sensors, aerial photographs and satellite imagery, as well as special software systems based on GIS. RSD are used for a more accurate evaluation of the seeding density, calculation of application rates and crop protection, more accurate prediction of yield and financial planning. Also, it must take into account local peculiarities of soil and climatic conditions. In some cases it may make it easier to determine the reasons for the deterioration of vegetation.

There are a number of systems for precision agriculture tasks: The Monitoring of Agriculture with Remote Sensing (MARS), Variogram Estimation and Spatial Prediction plus Error (VESPER), Ag Leader Insight etc. These systems are based on remote sensing processing methods, which allow effective detecting field areas that are infected by plant diseases. Detection and recognition of an infection on early stages of its development reduces costs of plant protective measures.

1. Functions and structure of the complex

The main task of developed hardware and software complex is preparing of morbidity maps of agricultural vegetation (potato) for fertilizer application for healing and prevention of plant diseases.

A series of algorithms for additional feature extraction and processing for agricultural fields images of different spatial resolution are developed by the author [7]. These algorithms can be used to create a technique of data processing which is used as basis for developing of decision making support system (DMSS) for the complex for precision farming tasks. The DMSS is used as a core of the developed complex.

To make decisions about state of the vegetation and amount of applied fertilizers, following steps should be performed:

1. Performing of pre-processing: filtering, white balance correction, data georeferencing.
2. Calculation of the additional features for each processed image: texture, fractal (fractal dimension) and color (ranges saturation and hue of different classes of objects, the normalized reduced histogram for training set of neural network classifier).
3. Performing of multi-criteria threshold and joint segmentation of color, texture and fractal features to detection of different areas of the original image: vegetation (affected and healthy), soil, vegetation and soil boundary, foreign objects.
4. Training of the proposed neural network classifier and recognition of aerial photographs to forming morbidity map which shows areas of soil, healthy and diseased plants and used as a basis for calculation of statistical indicators of productivity.
5. Forming of maps of morbidity rate which is input for control subsystem, which carries out application of plant protection products for their treatment.

The proposed hardware and software complex is designed for monitoring of beginning and development of diseases of vegetation as a part of GIS for precision farming. Structure of the complex is shown in fig. 1.

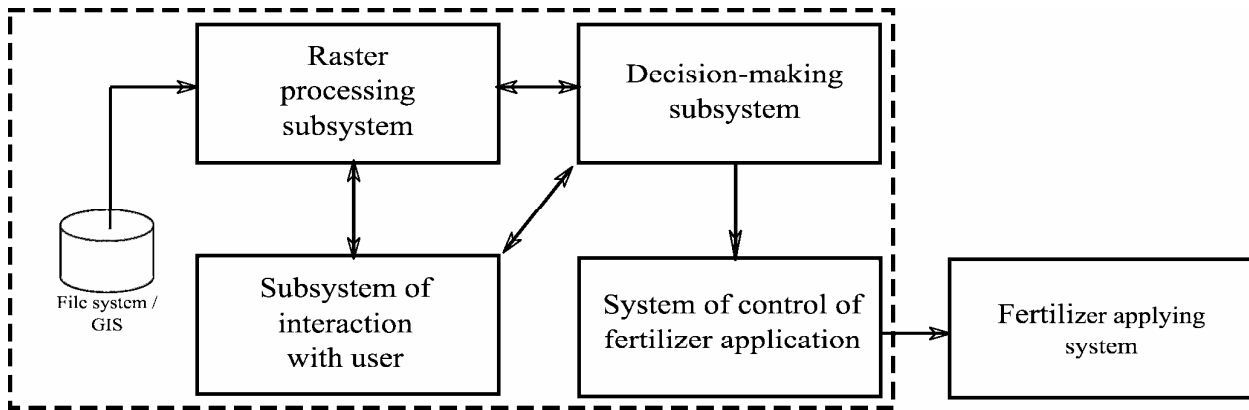


Fig. 1. The scheme hardware and software complex for precision farming tasks

2. Original Data

Agricultural field color images received with help of high resolution digital shooting are an object of our research (fig. 2). Lower value of this quantity equals higher spatial resolution of the image. In this article, if a side of a square is less than 0.6 cm, a spatial resolution is considered as high, otherwise – as low. We need to solve the problem of recognition for mapping of a disease. This can be done by recognizing the original image or by recognizing the received special area.

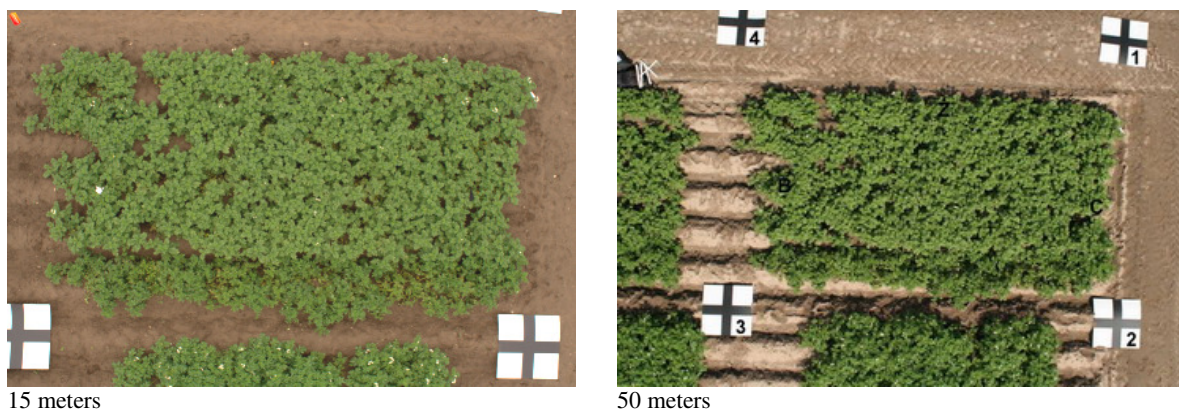


Fig. 2. Examples of original aerial photographs

3. Textures and Fractals

A texture is one of the major characteristics used for identification of objects or areas on the image [8]. The structures are subdivided into fine-grained, coarse-grained, smooth, granulated and undulating in according with used base attributes and interactions between them. In view of interaction degree of the base elements the structures are subdivided into strong (interaction submits to some rule) and weak (interaction has casual character).

An essence of the proposed method of textural characteristics calculation consists in calculation of separate channel images signatures with their subsequent association with use of factors which values depend on vegetation type and condition. An example of the obtained textures is shown in fig. 2. The example of textural characteristics calculation result is resulted on fig. 3, where visualization of calculated values Contrast is resulted. Contrast approximates 1 at a small variation of original data, and it vanishes at greater variation [9].

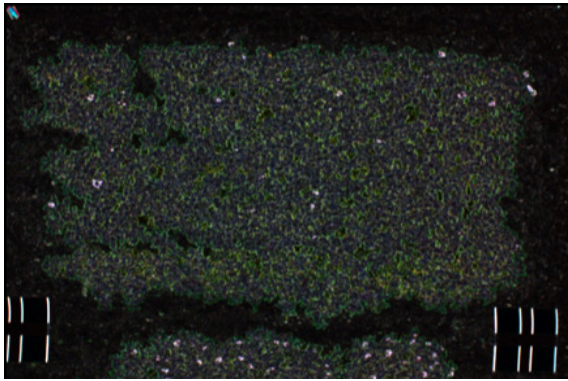


Fig. 3. Textural characteristics for the photograph shown in fig. 2 (15 meters)

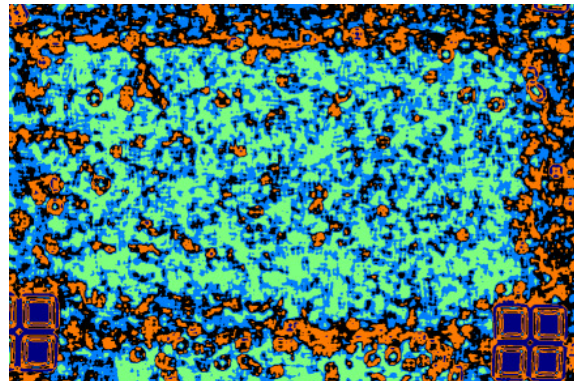


Fig. 4. Fractal signatures of various image areas

Fractal signatures calculation is based on the fact that quantified values of bidimensional signal intensity are located between two functions named the top and bottom surfaces [10]. Top surface U contains a set of points which values always exceed an intensity of the original signal. Bottom surface L has values of points which always are lower of the original image.

Results of fractal signatures calculation algorithm are presented in fig. 4 (visualization of the calculated values fractal signatures is resulted).

4. Joint Fuzzy Segmentation

Joint segmentation algorithm is based on co-processing of the source images and their fractal and textural features (i.e., an original color image is complemented by images of texture and fractal features).

Matrices of color features of original image, as well as the textural and fractal features computed for each color channel of the original image are used as the feature space.

Color ranges of corresponding healthy and diseased parts of potato fields obtained from an expert are used as color features. The algorithm is intended for segmentation of two-dimensional data representing matrices of various features of the original image such as color channels, textural and fractal features. Thus, the segmentation algorithm is executed in an N -

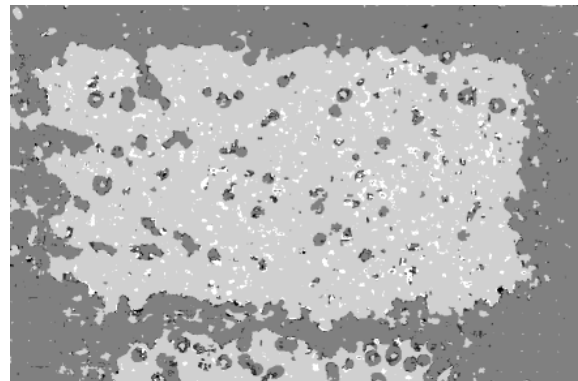


Fig. 5. Joint segmentation result example

dimensional space of attributes (where N is the number of characteristics used) where each dimension can be taken with a certain weight coefficient.

Obtained segmentation result (fig. 5) allows in an automatic mode to detect areas on which a disease is developed. The knowledge of an allocation of such areas allows determining requirement of fertilizers and other chemicals. It allows making agricultural works more effective and less expensive.

5. Color Features

Next step of processing is image recognition based on analysis of color features of various objects types. The analysis showed that within the same type, the features are differed slightly and are independent of spatial resolution. At the same time these features have some differences for different objects types. These differences in color features for each color channel (R , G , B or H , S , V) are offered to use for processing.

Color features of images are represented as normalized reduced histograms. A number of individual brightness elements is decreased and can be equal to zero for a small images. It creates distortions of a histogram. To reduce influence of the distortions it is proposed to use a histogram of intervals of brightness – the histogram based on a set of elements of brightness in each segment. Such histogram is

called a reduced. To ensure interoperability between the histograms of various sizes of images we use a normalization procedure.

6. Perceptron as a Classifier

To classify image areas a multilayer perceptron [11] is used with $N \times L$ inputs, where: N – number of segments of the normalized histogram, which is input of the perceptron, L – number of channels. One hidden layer contains 32×3 neurons (number of neurons is chosen experimentally), and an output layer containing three neurons corresponding to object types. A logistic activation function with sigmoid shape is used for each of the neurons. Data sample for classification is formed by a window of size $K \times K$ pixels scanning the original image.

Back-propagation algorithm is used to adjust weights of the perceptron. The normalized histograms of objects selected by operator are used as training sets. More than 1000 images are used for each class.

7. Evaluation of Quality Recognition

The results of evaluation of recognition based on described algorithms are shown in table.

Recognition algorithms evaluation

Algorithm	Error, %	Algorithm	Error, %
Fuzzy C-Means	18.7	Gustafson-Kessel	11.7
Gath-Geva	15.5	C-Means	15.9
SVM (HSV)	13.3	SVM (RGB)	16.7
Random Forest (HSV)	9.5	Random Forest (RGB)	15.4
Developed algorithm using HSV	4.9	Developed algorithm using RGB	4.8

Conclusion

The technique of recognition of vegetation state for decision support system for monitoring agricultural fields was proposed. The proposed algorithms for image segmentation, identification of specified areas and neural network classification are used as a core of the decision support system. These algorithms allow to build additional features and to configure algorithms for processing specific images in order to reduce time complexity and improve reliability of identification.

The structure of the complex was developed, that is based on the proposed technique of recognizing the state of vegetation from aerial photographs stored in GIS. The results of data processing can be used in the control system of fertilizers applying mechanisms for treatment and prevention of diseases of potato.

Scientific significance of these results consists in detection disease areas of agricultural plants fields which, which can be used in problems of precision farming. Practical importance consists of application of the developed complex for cultivation of green products at reduced cost. Possible area of application is remote sensing of the Earth (in precision farming, forestry).

References

1. Bonnefon, R. Geographic information system updating using remote sensing images / R. Bonnefon, P. Dherete, J. Desachy // *Pattern Recognition Letters*. – 2002. – Vol. 23. – P. 1073–1083.
2. Multi-sensor NDVI data continuity : Uncertainties and implications for vegetation monitoring applications / W.J.D. vanLeeuwen [et al.] // *Remote Sensing of Environment*. – 2006. – Vol. 3. – P. 67–81.
3. Rubtsov, S.A. Aerospace equipment and technologies for precision farming / S.A. Rubtsov, I.N. Golovanev. A.N. Kashtanov. – M., 2008. – 330 p. [In Russian]
4. Maselli, F. Evaluation of statistical methods to estimate forest volume in a Mediterranean region / F. Maselli, M. Chiesi // *IEEE Transactions on Geoscience and Remote Sensing*. – 2006. – Vol. 44, no. 8. – P. 2239–2250.
5. Ritchie, J.C. Remote sensing techniques to assess water quality / J.C. Ritchie, P.V. Zimba, J.H. Everitt // *Photogrammetric Engineering and Remote Sensing*. – 2003. – Vol. 69, no. 6. – P. 695–704.
6. Turner, M.G. Landscape ecology in theory and practice / M.G. Turner, R.H. Gardner, R.V. O'Neill. – Springer-Verlag, 2001. – 417 p.
7. Special Areas Detection and Recognition on Agricultural Fields Images / V. Ganchenko [et al.] // *Digital Image and Signal Processing for Measurement Systems* / ed. by Richard Duro, Fernando Pena. – River Publishers, 2012. – Ch. 8. – P. 201–233.

8. Starovoytov, V.V. Local geometrical methods of digital processing and the analysis of image / V.V. Starovoytov. – Minsk : IEC BAS, 1997. – 284 p.
9. Haralick, R.M. Textural Features for Image Classification / R.M. Haralick, K. Shanmugam, I. Dinstein // IEEE Transactions on Systems, Man and Cybernetics. – 1973. – No. 6. – P. 610–621.
10. Feder, E. Fractals / E. Feder. – M. : Mir, 1991. – 254 p. [In Russian]
11. Haykin, S. Neural networks: A Comprehensive Foundation / S. Haykin. – 2nd ed. – M. : Williams, 2006. – 1104 p. [In Russian]

USER-ORIENTED RECOGNITION OF INTELLIGENT INFORMATION OBJECTS IN DISTRIBUTED DYNAMIC INFORMATIONAL WEB-SPACE

A. Gladun¹, J. Rogushina², A. Andrushevich³, A. Kurbatski³

¹International Research and Training Center of Information Technologies and Systems of National Academy of Sciences Ukraine, Kiev;

²Institute of software systems of National Academy of Sciences Ukraine, Kiev;

³Software Engineering Department, Faculty of Applied Mathematics and Computer Science, Belarusian State University, Minsk
e-mail: glanat@yahoo.com

The idea of applying semantic web technologies to the area of recognition of information objects has resulted in a number of research activities and initiatives that have been recently developed. The feasibility of using ontologies to represent personalized, user-oriented and problem-oriented knowledge about intelligent information objects in distributed Web-based applications is proven. Methods of ontology use for recognition of information about complex structured information objects, interesting to user are proposed.

Introduction

Applications that work in distributed dynamic informational space for solving different user's problems, have access to a lot of solid multifarious inconsistent informational resources. Though, we try to recognize objects containing information useful for problem solving. Classification (class from user ontology) and definition (ontology class specification) of these objects depend on user tasks [1].

1. Ontology as a means of knowledge formalizing

The selection of knowledge representation formalisms for the majority of usual applications is determined only by problem to be solved and by different preferences of developers. But for intelligent Web-based applications this selection has its own peculiarity: considering their functioning in open information space they require the constant updating of knowledge from the external environment. Therefore it is advisable to use such interoperable knowledge representation as ontology which has been already established today in form of recognized standards, the representation languages, instrumental tools for editing and inference, as well as available fundamental mathematical basis.

At the same time, till now, to the best of our knowledge there are no accurately formulated technologies of knowledge management based on ontologies which could be directly implemented in applied systems, including the systems for recognition of certain information objects which are interesting for the user and necessary for the solving of user's problems.

The perception and recognition of information objects (IO) are the most important problems for working out of intelligent information systems which are based on knowledge [2]. Features of IO recognition in the distributed Web-based applications are their dynamism (for example, both standards and languages of the description of the metadata, and sources of data can vary), and dependence of distinguished IO structure from the peculiarity of problem to be solved.

Therefore, the creation of intelligent industrial systems based on ontologies, in environment of continuous organizational and technological changes requires methods and tools not only for ontology creation, but also for the whole complex of related problems - change management, estimations, personification, separation, mapping and integration etc.

Ontology mapping is the establishment of correspondences between two or more ontologies, and the integration of ontologies (ontology merging) - creation of a new ontology from multiple source ontologies. Ontology alignment, or ontology matching, is the process of determining correspondences between concepts. Ontology merging and alignment is one of the effective methods for ontology sharing and reuse on the Semantic Web. Ontology merging defines the act of bringing together two or more conceptually divergent ontologies or the instance data associated to this ontologies. This merging process can be performed in a number of ways: manually, semi-automatically or automatically.

2. Intelligent data objects in an ontological analysis

An *object* (from the point of view of programming and, in particular, the object-oriented approach) represents some entity in virtual space which possesses a certain condition and behavior, has preset values of properties (attributes) and operations over them (methods).

Typically, during the study of objects the belonging of each object to one or several classes is allocated. This belonging in turn causes the behavior of the object, i.e. is a model of this object [3].

The information object is an extension of the program object. The information object represents the certain entity comprising data in information system about any real or virtual object (subject, being, event, process etc.) that is the unique identified material or non-material entity of the real world which describes its structure, attributes, restriction of integrity and, probably, behavior.

For example, a person, a publication, an organization, a house, a building [4], a city can be an object, and the description of some object (information about some features of this object) can be an information object. For the Semantic Web concept and the Web information space such objects of the non-material world, as ontologies, software agents, Web-services, information resources, metadata, databases, etc. also belong to the information objects.

The other aspect of IO classification is a data type viewpoint. Now the Web proposes a lot of multimedia data for different applications and with different expressiveness. Informational Resources (IR) represented in the Internet can be classified on textual and multimedia ones, static and dynamic, structured and not structured etc., but every IR has some semantics and is concerned with some subject domain. IRs can be also seen as IOs. In the process of information retrieval [5] it is very important to discover IR concerned with the domain interested to user. Structures textual information in the Internet is mainly given in HTML and XML formats. The subject domain of textual IR can be define by two ways: 1) analyzing of IR textual content and 2) considering metadata of these IR. There is a great deal of the widespread formats for a storing audio and video information, 3D-scripts and images. The multimedia resources are accessible for indexation much worse than textual information. Therefore for multimedia IR only the second way is efficient. Metadata contains machine-readable information about the document that can be automatically processed by computer. Now the most perspective and common metadata model is RDF (Resource Description Framework) based on XML. Though, the ontology of the Web multimedia resources can help users in the IR type description for problem solving in correspondence with software for their processing (fig. 1).

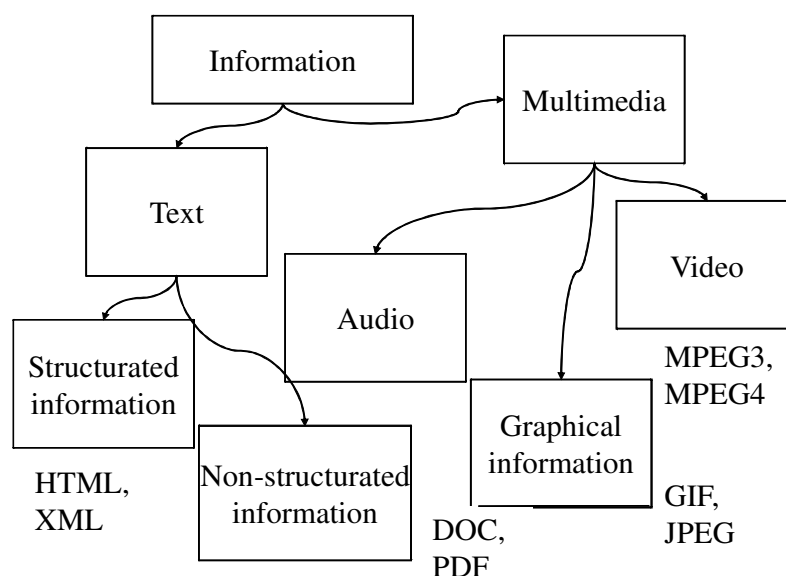


Fig. 1. Types of multimedia informational resources of the Web

The main contributions of this work are three folds:

1. Proposing a new classification mechanism for dynamic multimedia objects using users' access history.

2. Comparing experimentally an approach that analyzes each user's accesses against one considering collective user behavior only, and demonstrating a much higher classification quality with the former approach [6].

3. Investigating the relation of four factors to classification quality, and evaluating four corresponding rules for filtering classification results.

Classification techniques can be categorized as manual or automatic. Manual classification can deal with different types of data and has high accuracy. A pioneering search engine Yahoo manually indexes its contents. The two major auction sites Amazon and eBay require sellers to classify the auctions they upload to the sites. The drawback of manual classification is the inefficiency in dealing with a large number of objects in a complicated category structure.

Automatic classification systems are either rule-based or machine-learning systems; and machine-learning systems are further divided into supervised (classification) or unsupervised (clustering) [3, 6]. According to the knowledge on which the classification is based, automatic classifications can also be identified as content-based or link structure-based. Beside documents, the objects to be classified can also be search engines or sites on specific topics. The general model of multimedia IRs is proposed on fig. 2.

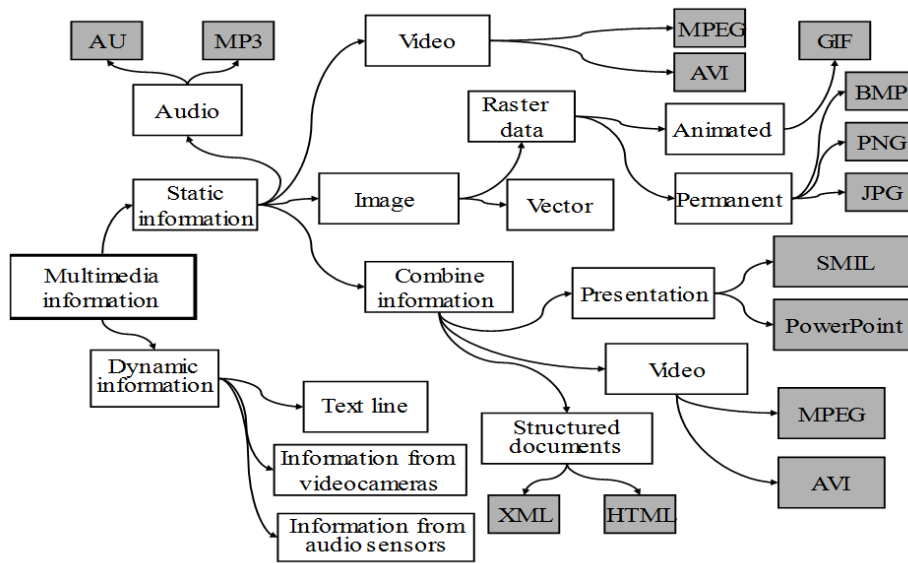


Fig. 2. General model of multimedia IRs

Nowadays, a lot of Web applications are intelligent and use knowledge about some subject domain or produce some new knowledge. In such applications knowledge is represented in interoperable form and can be reusable. For such representation ontological approach is widely used because ontologies have a fundamental theoretical foundation (descriptive logic). That's why in the modern Web the significant part of content is intelligent and knowledge-oriented – ontologies, semantic Wikipedia items, elements of semantic markup, semantic metadata about other resources etc (fig. 3). At the present stage of IT in the majority of cases intelligent Web applications use standards and technologies of knowledge management developed by Semantic Web project. Ontologies are an important building block of knowledge in the Semantic Web. They provide a shared and common understanding of a domain that can be communicated across people and applications. An important advantage of the Semantic Web is that people can collaboratively create ontologies and build common vocabulary without centralized control. One building block of Semantic Web ontologies is a Semantic Web Term (SWT) which plays the role of a word in natural languages. The set of SWT is a bridge between the RDF statements with formal semantics defined in RDF(S) and OWL. The social Web provides the knowledge about persons and communities. Ontological approach enables to link such information with some classes of ontology that are interesting to user. This process of class recognition of heterogeneous data is a kind of pattern recognition.

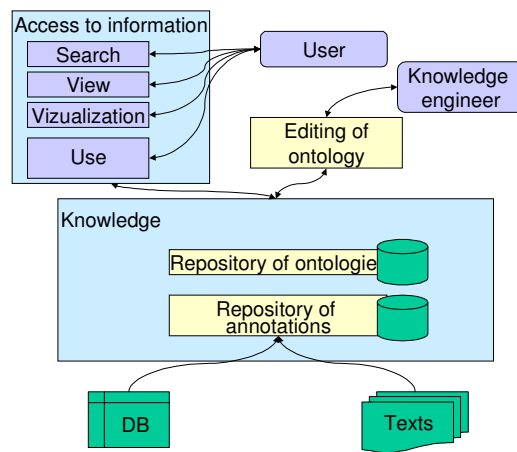


Fig. 3. Main elements of ontological knowledge

Wikipedia defines *pattern recognition* as the assignment of the source data to a certain class by means of allocation of the significant features that characterize these data from the total weight of non-essential data. An example of pattern recognition is classification, which attempts to assign each input value to one of a given set of classes (for example, determine whether a given email is “spam” or “non-spam”). However, pattern recognition is a more general problem that encompasses other types of output as well. Other examples are regression, which assigns a real-valued output to each input; sequence labeling, which assigns a class to each member of a sequence of values (for example, part of speech tagging, which assigns a part of speech to each word in an input sentence); and parsing, which assigns a parse tree to an input sentence, describing the syntactic structure of the sentence.

The classic statement of the problem of *pattern recognition* [7, 8]: by the set of given objects it is necessary to classify them. The set of objects consists of subsets called classes. Information about the classes, description of the whole set, information about some classified objects and information about an object, which belonging to some class is unknown but interesting to user are specified. Under the available information about classes and the object description the class is required to establish.

In modern intelligent applications this problem is usually transformed as follows - it is necessary to retrieve from the available (via Web) or derived in any other way relevant to the task informational resources (IR) the information relating to certain concepts (classes or class instances) and relations among them (their structure is reflected by domain ontology).

For example, the user needs to find a set of executors to perform a specific research project. Domain ontology allows to know that the desired object is an entity belonging to the class of “human” and having parameters such as education, experience, skills, availability of publications, degree, diplomas and work experience, etc.

But only the availability of relevant ontology allows the *problem-specific structuration* of information extracted from the IRs, because information used during the solving of another problem and extracted from the same IRs about the same objects may be absolutely different (for example, information for detection of effective ways of treatment of patients with similar symptoms differs from the information extracted from the same IRs about the same persons for finding dog owners with similar breeds).

Problem of perception, recognition and interpretation of objects is a complex task which is divided into separate subtasks [8,9]. Thus traditional pattern recognition, speech recognition and text recognition are only the special cases of a considerably more general problem.

Recognition of information object (IO) involves detection of information interesting to user about a particular IO in some IR content. For example, face recognition is a detection of elements that characterize the look data of IO of type "human" in the graphic IR [9].

3. Use of ontologies for structurization of recognised data

Ontology can be considered as a basis for IO structure representation, i.e. IO is an ontology class, and various IRs are the sources for creating of class instances with interesting to user data. This approach allows to integrate information coming from different sources, and to generate the necessary to user knowledge. The task is subdivided into several sub-tasks:

- *Creation (or search) of ontology*, representing the structure of the IO (or of the set of the IO), knowledge about which is necessary to solving of the user's problem;
- *Retrieval of IRs* that explicitly or implicitly contain data about these IOs;
- *Acquisition of knowledge* about IO from IRs;
- *Representation of the retrieved knowledge* in the comprehensible and convenient to the user form.

General architecture of ontology-based recognition process links user problems with domain ontologies that are used in the recognition process of different IOs that are interested to users (fig. 4).

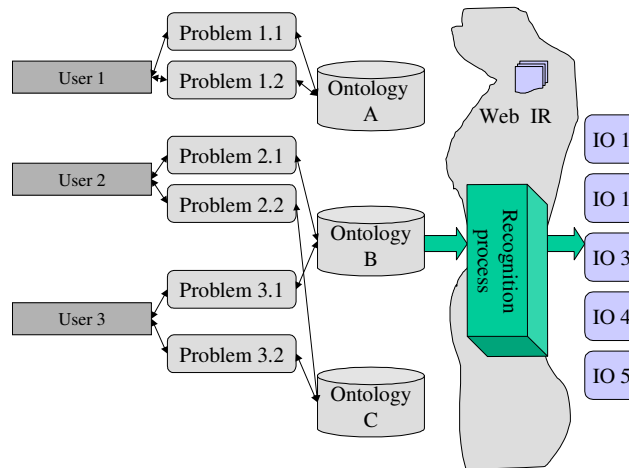


Fig. 4. General architecture of ontology-based recognition process

It should be noted that in many cases this process is iterative, and during the problem resolving the information about IO has to be updated by extracting the relevant information from those IR which are available to the user (for example, via the Web or corporate network).

Conclusion

Content analysis and link structure analysis are the two dominant approaches for document classification. As observed by many in the literature [8], a major shortcoming of content analysis is that it usually requires text and so cannot be directly applied to classifying multimedia information. Link structure analysis, on the other hand, requires that changes be reflected promptly in the links around a dynamic object. Neither of the two techniques guarantees automatic and prompt classification for multimedia objects that may be changing quickly.

Most sites with dynamic information objects are, however, stable in their category structure. In addition, popular sites accumulate a large amount of information about their users' accesses to the sites. Both the structural stability and the rich information about users' access history are valuable resources but their utilization in content classification is not yet investigated sufficiently.

This work proposes a framework to classify an object into a category structure by analyzing users' traversals in the category structure. Use of ontologies as a source of information about the structuring of recognized IO is caused by properties of ontologies and their role in today's distributed intelligent applications. Furthermore, ontologies themselves can be updated with new knowledge as a result of the recognition process.

References

1. Kirasic, D. *Ontology-Based Design Pattern Recognition Knowledge-Based Intelligent Information and Engineering Systems / D. Kirasic, D. Basch // Lecture Notes in Computer Science. – 2008. – Vol. 5177. – P. 384–393.*
2. Kikumi, K. *Taisuoka Architecture of Knowledge Structures and Cognitive Diagnosis: A Statistical Pattern Recognition and Classification Approach / K. Kikumi // Book Cognitively diagnostic assessment. – University of Iowa USA, 2003. – P. 327–361.*

3. Dumais, S. Hierarchical Classification of Web Content / S. Dumais, H. Chen // Proceedings of SIGIR. – 2000.
4. Towards Semantic Buildings: Goal-driven approach for building automation service allocation and control / A. Andrushevich [et al.] // Proc. of 15th IEEE Intern. Conf. on Emerging Technologies and Factory Automation (ETFFA 2010), September 13, 2010, Bilbao, Spain. – Bilbao, 2010.
5. Prometheus Framework for Fuzzy Information Retrieval in Semantic Spaces / A. Andrushevich [et al.] // Human – Computer Systems Interaction: Backgrounds and Applications / eds. Zd.S. Hippe, Ju.L. Kulikowski, T. Mroczek. – 2. ed. – Vol. Advances in Intelligent and Soft Computing. – Springer, 2011.
6. Chen, M. Categorizing Information Objects from User Access Patterns / M. Chen, A. LaPaugh, J. Pal Singh // Proceedings of CIKM'02. – McLean, Virginia, USA, 2002.
7. Rogushina, J. Management of Ontologic Knowledge for Recognition of Information Objects in the Distributed Recommending Systems / J. Rogushina, A. Gladun // Proceedings 3th Intern. Conf. Knowledge Management and Business Intelligence. – Kharkiv, 2013. – Vol. 1. – P. 78–82. [In Russian].
8. Rogushina, J. Ontology-based Competency Analyses in New Research Domains / J. Rogushina, A. Gladun // International Journal of Computing and Information Technology – CIT. – 2013. – Vol. 20, № 4. – P. 277–291.
9. Gladun, A. Cognitive Networks and Ontologic Knowledge for Increase of Adaptability and Quality of Service in Heterogeneous Wireless Environment / A. Gladun, J. Rogushina // Proceedings OSTIS-2012. – Minsk, 2012. – P. 493–500. [In Russian]

SKELETONIZATION ALGORITHM OF HIGH RESOLUTION VASCULAR DATA

D. Hancharou¹, A. Nedzved², S. Ablameyko¹
¹Belarusian State University, Minsk;

²United Institute of Informatics Problems of the NAS of Belarus, Minsk
e-mail: goncharovda@email.com

Analysis of vascular structures can greatly improve accuracy of morphometric studies that may be used for early medical diagnosis. Laws and processes in vascular tree can be described by skeletal structure. Currently, there are numerous skeletonization methods that focus on computational efficiency. However, when data are too large to be loaded into the memory of a personal computer, such approaches are difficult to use. This paper presents an algorithm for constructing of skeleton, topologically equivalent to the original object, which processes data locally and at the same time preserves global attribute of the object - the centeredness. Built skeleton is based on distance maps and then can be used to calculate the morphological and topological characteristics.

Introduction

Non-invasive tests allow to obtain diagnostic information excluding the possibility of entering viruses and bacteria into a body and relieving a patient from complex and unpleasant pain. At the same time, the tools required not only for carrying out such studies, but for analysis too. Skeleton is a convenient form for the analysis of objects, particularly for objects with complex topological structure. Comparison of features of vascular structures in norm and pathology allows us to understand many morphological processes.

Currently there are many algorithms for skeleton constructing [1], but an input image has to be processed entirely in the vast majority of them. However, modern medical imaging algorithms allow to get the images of high accuracy, therefore it leads to an increase in processed data. E.g., the analyzed images of vascular structures may occupy several gigabytes. Such image may not always be loaded into memory and special methods should be offered for images of this size.

The main difficulty in the skeleton construction is to preserve the global characteristics of the skeleton (e.g. centeredness) during local transformations. In this article, we propose generalized skeletonization algorithm. We also give the analysis of obtained results, the advantages and drawbacks of our approach.

1. Properties of vessels image

All vessels are divided into three category by section: longitudinal, transverse, mixed. Elongated vessels are characterized by elongated shape and can be divided to two classes: thin, thick. These classes have different algorithms of extraction and description. Transverse vessels have ellipse-like shape and can to have a few extraction algorithms. These algorithms depend on different condition of vessels representation. Pattern of vessels can be defined by: border, contour of cell, free space. This condition depends on type of histology sample. Therefore definition of vessels extraction method is choosing interactively by specialists (fig. 1).

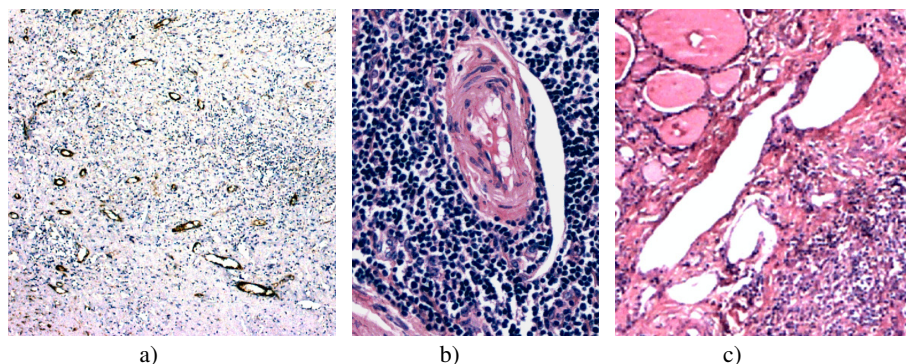


Fig. 1. Sample of vessels: a) transverse and mixed vessels that define by border; b) transverse that defined by contour of cell; c) longitudinal vessels that define by free space

2. Proposed algorithm

2.1. General Scheme

The general scheme of the proposed algorithm can be viewed in fig. 2.

On the first stage we acquire high resolution images with vessels. It is a mosaic with about a hundred of images, each of them contains about 2000x1000 voxels (15 Mb).

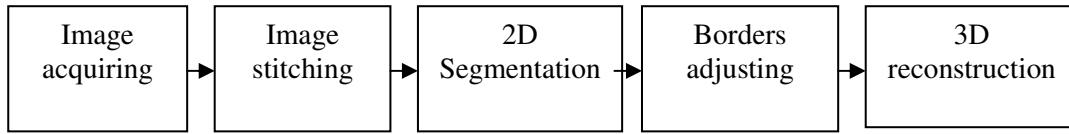


Fig. 2. General scheme of the algorithm

Obtained images have an overlap of about 50 voxel on each border between the mosaic images. For re-estimation of images positions stitch image stitching algorithm based on normalized statistical moment [2] can be applied. It is chosen because it optimized for high resolution images and robust to small rotations between subimages.

On the third stage for every subimage we apply the algorithm of color image processing of a histological image sample for membrane analysis described in [3]. Result of this operation is shown in fig. 3.

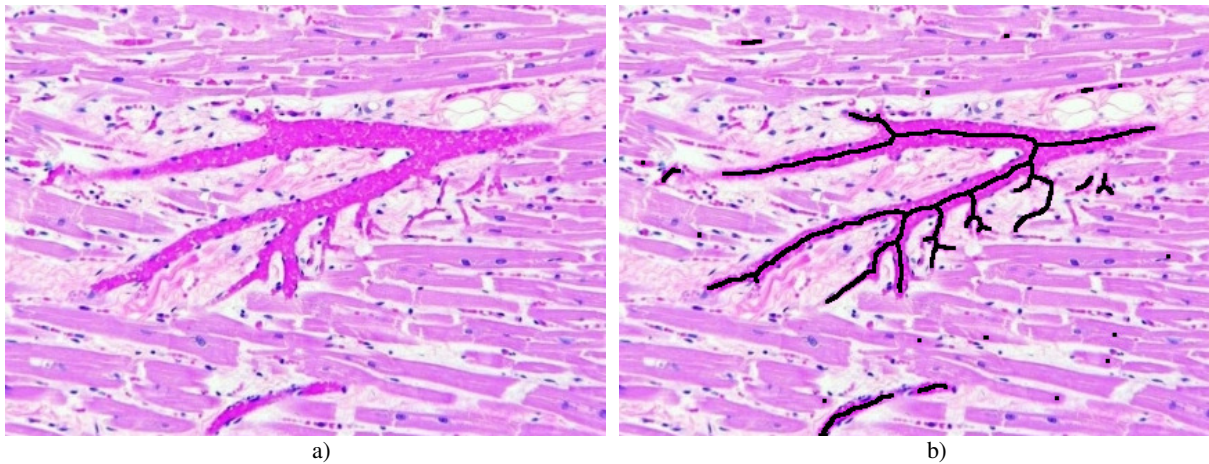


Fig. 3. Original histological image (a); extracted skeletons (b)

We will take a look more closely at two final steps.

2.2. Preserving centeredness

Let's consider two neighboring subimages (fig. 4, a). If the initial algorithm is not changed, some artifacts will appear around the borders (fig. 4, b) and obtained skeleton will not be centered for the object.

To resolve this problem, we remove only those points for which the distance to the boundary of the object is less than the distance to the border of subimage:

$$P_{deletable} = P_{deletable} \wedge dist_p < \min(d_x, d_y, d_z),$$

where $dist_p$ – distance map value for P voxel, d_x, d_y, d_z – the distance to the nearest point of the boundary subimages respectively for axes Ox, Oy, Oz .

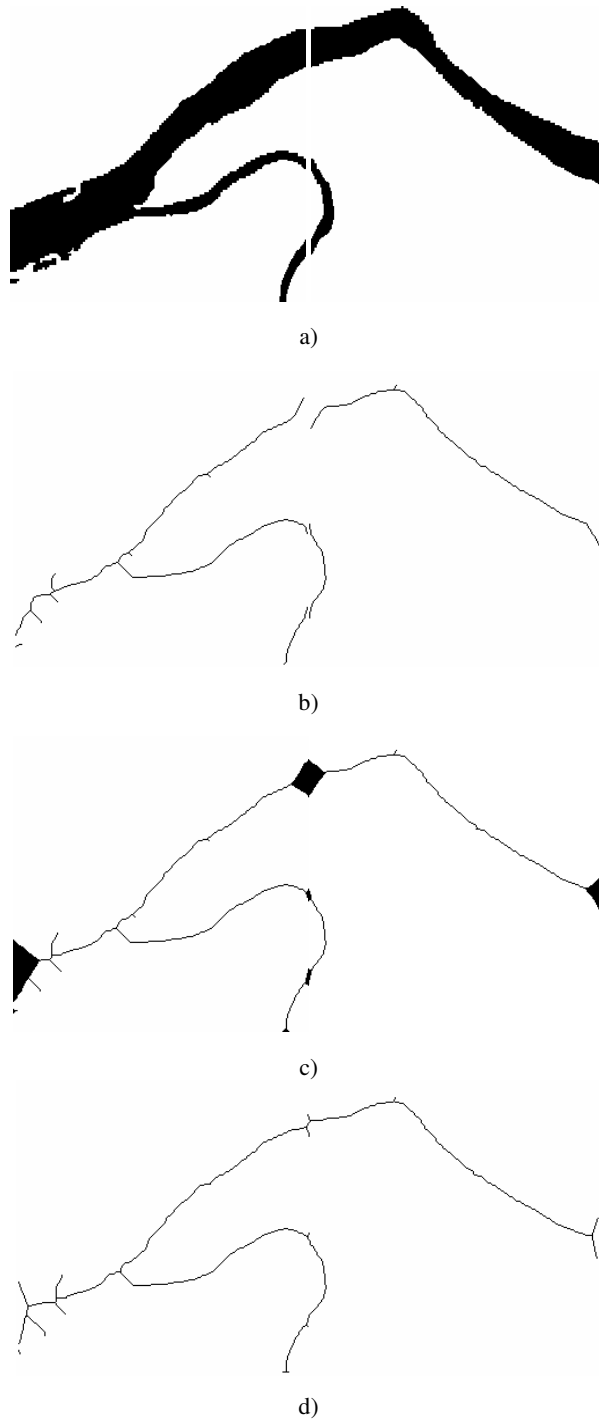


Fig. 4. Two neighboring subimages (a); result of stage 3 without correction (b); result of the first pass of the stage 4 (c); corrected skeleton (d)

This condition ensures that central points of the object will not be removed during this pass on the fourth stage (fig. 4, c).

This step is independent for each image and can be performed in parallel.

Unprocessed voxels are removed in the next steps of the stage. Their idea is that processing only occurs near the boundaries of the subimage. The subimage is cut into “strips” with width equal to doubled maximum distance on the distance map. This allows covering all voxels. Then, for each such “strip” used a slightly modified first step: in the removal condition is ignored for considered axis. For example, for the “strip” along the axis Ox condition removal is as follows:

$$P_{deletable} = P_{deletable} \wedge dist_p < \min(d_y, d_z).$$

Thus, the skeleton is thinned along each of the axes. After this stage we should to reconstruct skeleton for 3D-case.

3. Extension to 3D

The construction of intermediate layers can be realized by analysis of distance map properties [4]. Rides of distance map correspond to contours of optimal intermediate layer. Also they correspond to watershed lines. Taking in consideration aforesaid, we used technique [5] in order to reconstruct 3D-object surface from several closed, in general, non-planar curves.

Before that we applied dilatation to constructed skeleton. Detailed description of algorithm can found at [4], short scheme is shown on fig. 5, results on fig. 6.

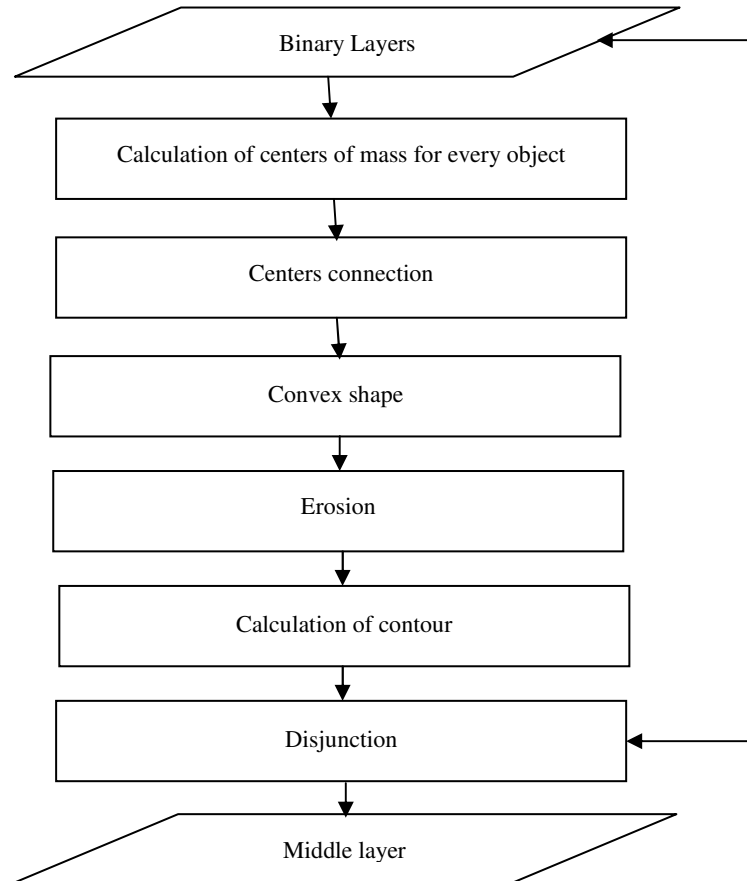


Fig. 5. Algorithm of construction of regions for definition

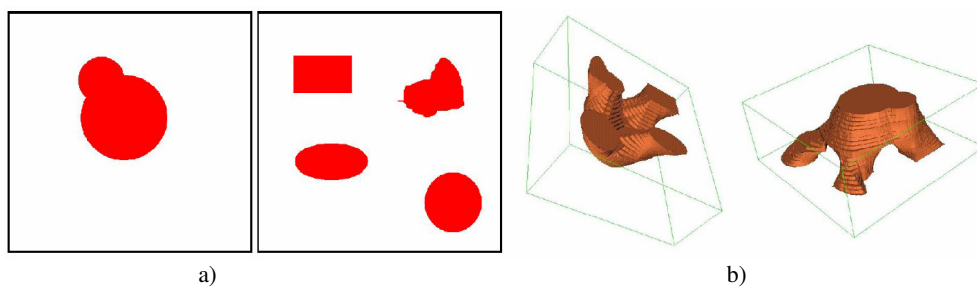


Fig. 6. Neighboring layers (a); reconstructed image (b)

Conclusion

The proposed algorithm has several advantages. Firstly, it allows processing a large amount of images, which is especially important for histology images as well as magnetic resonance images. Secondly, the considered algorithm can be easily parallelized, which will increase the speed of image processing and, consequently, the medical diagnosis. Thirdly, constructed skeleton contains information about the distance to the boundary of the object, and it can be used for further analysis.

The algorithm was tested on synthetic images and images of complex medical objects. It demonstrated good results for all cases. Compared to analogs, the proposed algorithm is self-contained – it works with raster layered images without additional transformations. In addition, the proposed algorithm is flexible; it has clearly marked stages, which can optionally be substituted with other alternatives.

References

1. Cornea, N. Curve-Skeleton Properties, Applications, and Algorithms / N. Cornea, D. Silver, P. Min // IEEE Transactions on Visualization and Computer Graphics. – 2007. – Vol. 13, № 3. – P. 530–548.
2. HANCHAROU, D.A. HISTOLOGICAL IMAGE STITCHING ALGORITHM BASED ON NORMALIZED STATISTICAL MOMENT / D.A. HANCHAROU, A.M. NEDZVED, V.V. STAROVOITOV // PROC. OF THE ELEVENTH INTERN. CONF. ON PATTERN RECOGNITION AND INFORMATION PROCESSING. – MINSK, 2011. – P. 200–203.
3. NEDZVED, A.M. EXTRACTION OF THIN COLOR PATTERN FROM IMAGES FOR HISTOLOGY INVESTIGATION / A.M. NEDZVED, V.V. STAROVOITOV // PROC. OF 2010 IEEE WORLD CONGRESS ON COMPUTATIONAL INTELLIGENCE. – BARCELONA, 2010. – P. 1922–1928.
4. HANCHAROU, D. 3D DISTANCE TRANSFORM AND ITS APPLICATION FOR PROCESSING OF MEDICAL IMAGES / D.A. HANCHAROU, A.M. NEDZVED, P.V. LUKASHEVICH // JOURNAL OF INFORMATION, CONTROL AND MANAGEMENT SYSTEMS. – 2010. – VOL. 8 (2). – P. 43–53.
5. NEDZVED, A.M. RESTORING FORM THE BULK OF MEDICAL FACILITIES BY REMOTE MAPPING OF CROSS SECTIONS / A.M. NEDZVED, P.V. LUKASHEVICH, D.A. GONCHAROV // PROC. OF MATERIALS OF THE INTERNATIONAL SCIENTIFIC CONFERENCE “ARTIFICIAL INTELLIGENCE. INTELLECTUAL SYSTEMS”. – KATSIVELI, 2010. – P. 324–327.

THE SYSTEM OF GENERATION OF PHONETIC TRANSCRIPTIONS FOR INPUT ELECTRONIC TEXTS IN BELARUSIAN

Yu. Hetsevich, V. Mandzik, V. Rusak, A. Hiuntar, T. Okrut, B. Lobanov, S. Lysy,
Dz. Dzenisiuk,

United Institute of Informatics Problems of the NAS of Belarus, Minsk
e-mail: yury.hetsevich@gmail.com

The paper discusses the process of creating a system which automatically produces phonetic transcriptions for input electronic texts in the Belarusian language according to the Cyrillic and Latin alphabetic systems of phonetic notation. The general scheme, the working algorithm, the structure and contents of the specialized knowledge base with correspondences “an allophone – a transcription symbol” are described.

Introduction

The authors' aim is to create a system of the automatic generation of phonetic transcriptions (GT) for electronic texts in the Belarusian language. Fields of use may include various services and applications, such as online text-to-speech synthesis [1], natural-speech interfaces [2] – in order to define the characteristics (i.e. peculiarities of a representation of the sounds of oral language), according to which any Belarusian word can get the correct automatic pronunciation. At present the problem is urgent because there are no specialized algorithms for Belarusian words to automatically receive the Cyrillic or Latin transcriptions.

In order to achieve our goals, we have established the following specific objectives:

- to develop a knowledge base with correspondences “an allophone – a transcription symbol”;
- to implement program algorithms which transcribe electronic texts (i.e. generate phonetic notation of each word) according to the Cyrillic and Latin phonetic alphabets;
- on the basis of a GT-system, to create a GT-service accessible via the Internet;
- to work out a generalized scheme of a GT-system for other languages.

For the construction of a knowledge base for the Cyrillic transcription of words, language experts have used scientific materials on the theory of the Belarusian language [3, 4]. The research work [5] has become the basis for a knowledge base of the first Latin transcription, and the resources [6, 7] have been used for a knowledge base of the International Phonetic Alphabet (IPA) which is the other phonetic notation system based primarily on the Latin alphabet.

1. The algorithm for the automatic generation of phonetic transcriptions

In order to build the system which produces phonetic transcriptions, we assume that a user puts an electronic, orthographically correct text in the allophonic representation at the input of the system (fig. 1). Further, the system is supposed to produce transcriptions according to several phonetic alphabets at the output. The allophonic representation of orthographically correct texts can be obtained with the help of an online text-to-speech synthesizer available at www.corpus.by/tts3 [8].

Let us take the following orthographic text with marked stresses, and phonemic words as an example:

*Мо=й ро+дны ку+т, я=к ты= мне= ми+лы!
Забы+ць цябе= няЪма+ю си+лы!*

After this text is processed by the text-to-speech synthesizer, its allophonic representation will be as follows:

*M004,O113,J'013,/,R032,O022,D001,N004,Y322,/,K001,U032,T000,/,#C3,
J'002,A142,K004,/,T002,Y121,/,M001,N'004,E143,/,M'002,I041,L004,Y310,/,#E2,*

Once an electronic text in the allophonic representation is at the input of the GT-system, it becomes possible for a user to select a language of the input text, and a desirable phonetic transcription. For now the system can convert Belarusian electronic texts to their phonetic transcription using three types of phonetic notation systems: transcription by Cyrillic characters (let us sign it with Tr_1), transcription by Latin characters on the basis of [5] (Tr_2), and transcription by Latin characters according to the international standardized format, namely IPA (Tr_3).

The results of using the algorithm are given to a user in the form of a transcribed text according to the desirable phonetic alphabet. At the same time linguistic experts and software engineers receive all the input, output, and analytic data (a size of a text, an IP-address, time, etc.) for efficient correction of errors in the process of producing transcriptions.

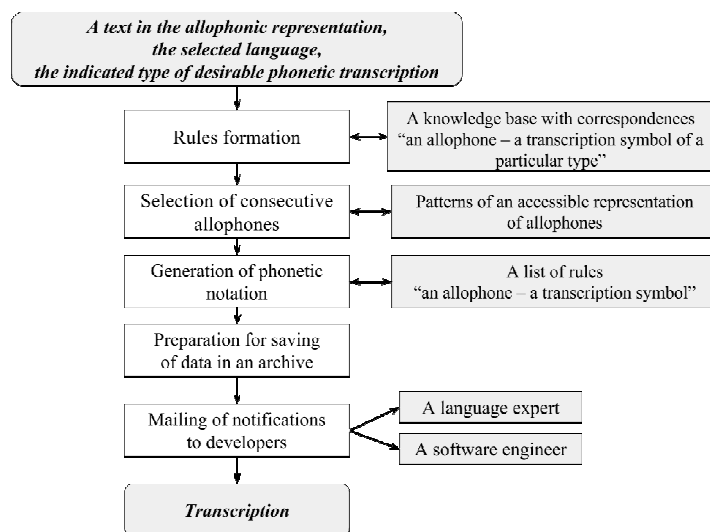


Fig. 1. A generalized scheme of the algorithm for the automatic generation of phonetic transcriptions

Let us consider in detail the algorithm of work of the GT-system.

The algorithm requires the following data at the input: an electronic text T in the allophonic representation, a language L , and a desirable type of a phonetic alphabet according to which the transcription process will be performed Tr_i . Further the algorithm fulfills the following steps:

Step 1. *Formation of replacement rules.* At first the algorithm accesses a knowledge base of the language L chosen by a user. The knowledge base contains correspondences “an allophone – a transcription symbol”. Then a number of replacement rules $R = \langle R_1, \dots, R_n \rangle$ are formed, where $R_i = \langle a_i, tr_{i1}, tr_{i2}, tr_{i3} \rangle$; a_i is an allophone code; tr_{i1} , tr_{i2} , tr_{i3} denote characters of three types of transcriptions of an allophone a_i ; $I = 1 \dots n$; n is the quantity of rules.

Step 2. *Allophonic representation of a text.* In the text T , allophones are selected according to the pattern Pt of their accessible view. The pattern Pt can be formalized by means of the language of regular expressions in the following way: $Pt = [A-Z]\{1,3\}[\prime]\{0,1\}[\backslash d]\{3\}$, where $[A-Z]\{1,3\}$ is one, two or three capital letters of the Latin alphabet, $[\prime]\{0,1\}$ means a possible apostrophe sign, $[\backslash d]\{3\}$ signify three digits $(0, \dots, 9)$. The fulfillment of this step is necessary because theoretically a user can place a text in any representation at the input of the system. In this case conversion of such a text requires its preliminary normalization, i.e. extraction of a text in the allophonic representation T_a out of the input text T .

Step 3. *Generation of transcriptions, namely phonetic notation.* For each allophone a_i which is sequentially taken from the text T_a , a transcription symbol of one or another type tr_{ij} is generated according to rules R which have been created earlier. The consecutive processing of the whole text T_a results in the final transcription Tr of the selected type Tr_i .

Step 4. *Preparation for data saving.* During the automated phonetic notation, all the information on the process is gathered. This information includes various input (an allophonic text T , a language L , a desirable type of transcription Tr_i), output (a transcribed text Tr), and analytic (a size of a text, an IP-address, request time, etc.) data. Afterwards the data collection is used by developers to efficiently correct errors (if they occur) in the process of transcriptions generation. It should be noted, that if one or another element of the knowledge base “an allophone – a transcription symbol” is missing, the system gathers this information as well.

Step 5. *Mailing of notifications to developers.* After the fulfillment of step 4 the above-mentioned data are attached to an e-mail letter which is sent to developers – a language expert and a software engineer – for determining the type of a problematic situation (linguistic mistakes, errors of algorithms and program codes), and the best way to solve it.

Step 6. *The end of work of the algorithm.* As a result, after the algorithm is accomplished a user receives a transcribed text. Fig. 2 illustrates some examples of phonetically transcribed text lines.

Transcription in cyrillic letters:	Transcription in latin letters:	Transcription in International Phonetic Alphabet :
[m'oi] [r'odny] [k'ut] [j'ak] [t'y] [m'n'e] [m'ily] [zab'yc'] [c'ab'e] [n'amaj'u] [s'ily]	[mòj] [ródný] [kút] [jàk] [tỳ] [mn'è] [m'ily] [zabýts'] [tš'ab'è] [n'amáj'u] [s'ily]	[m,oj] [r'odni] [k'ut] [jæk] [tɨ] [mn'e] [m'ili] [zɛb'ɨtɕ] [tʂɛb'ɛ] [n'ɛm'eju] [s'ili]
a)	b)	c)

Fig. 2. The example of the resulting text which has been transcribed according to the Cyrillic alphabet (a), simplified Latin alphabet (b), and IPA (c)

High-quality work of the algorithm of the GT-system requires the existence of rules R which can give the corresponding transcriptions $tr_{i1}, tr_{i2}, tr_{i3}$ to any input allophone. A linguistic expert was instructed to develop records for a knowledge base which would correspond to any possible input set of allophones of Belarusian texts according to [9]. In addition each record of a knowledge base is required to be actualized (i.e. filled with all types of transcriptions). Let us consider the process of developing a knowledge base for the system of transcriptions generation.

2. The development of a knowledge base for the system of transcriptions generation

For automatic phonetic notation, it is necessary to create tables of correspondences of allophones with transcription symbols. Initially, our work was carried out in the direction of creating a generator of the Cyrillic transcription. That is why at first we started to create a list of “allophone – transcription” correspondences for this type of transcription. For this purpose, an expert selected 720 words in order to cover all possible occurrences of allophones by the text-to-speech synthesizer. Further a specialist in phonetics used these words for the creation of a knowledge base for the Cyrillic transcription and its analysis (table 1). It was stated that in order to correctly and unambiguously transcribe a word (in its allophonic representation), the first three characters of each allophone are enough, instead of a maximum of five characters that might be given out by a phonetic processor of a text-to-speech synthesizer.

Table 1
Transformation of words into the Cyrillic transcription (an excerpt)

Word	Cyrillic transcription	Allophonic sequence
звѣрабo+й.	з'в'эрабoй.	Z'0,V'0,E2,R0,A2,B0,O0,J'0,
бамбѣ+жка.	бамб'oшка.	B0,A2,M0,B'0,O0,SH0,K0,A2,
адду+шына.	ад:ушына.	A2,D1,U0,SH0,Y3,N0,A2,

In the Cyrillic transcription, the following diacritics are used: the acute accent (´) and the grave accent (`) – the main and side verbal accents which are put directly above a stressed vowel; the colon (:) – a reduplication sign; and the apostrophe (´) – a palatalization sign.

After that, an expert correlated all existing allophones with the Cyrillic transcription characters for further automation of the process of transforming sequences of allophones into transcribed records according to the Cyrillic phonetic alphabet (table 2).

Table 2
Allophones with their corresponding symbols of the Cyrillic transcription, and examples of words containing them (an excerpt)

Allophone	Cyrillic transcription symbol	Word example
A0	á	катала+жка
A1	à	дзя=ржбюджэ+це
A2	a	ураджэ+нка.
A3	a	навако+лле.

The next step became the development of a knowledge base for the Latin transcription on the basis of the work by U. Koščanka [5]. Thus, the knowledge base with the characters of the Cyrillic alphabet was supplemented with their Latin equivalents and examples of Latin transcriptions with accents (table 3).

Table 3
Transformation of words into the Latin transcription (according to the [5]) (an excerpt)

Allophone	Word example (1)	Latin transcription symbol [5]	Word example (2)
E0	канстэ+бль.	é	kanstébl□
E1	тэ=леперада+ча.	è	tèl'ep'eradát□□a
E2	тэа+гр.	e	teátr

However, after we analyzed the results of the created table, it became obvious that the Latin transcription we had received could not be called international because it did not correspond to the International Phonetic Alphabet (IPA) [6, 7]. In order to achieve our goal we started to deeply analyze the IPA, and to search for equivalents of Belarusian sounds.

There are twenty eight different characters for vowel sounds in the IPA. These characters are systematized in a certain articulation classification [6, 7]. In order to identify relevant sounds, we compared this classification to the corresponding classification of vowels of the Belarusian language [3]. As a result, six necessary sounds were added. Next we found equivalents for all consonants.

Then two more columns were added to the described knowledge base. One of them illustrates sounds in the representation according to the international format, while the second one contains examples of words converted into the international transcription (table 4). At the same time stresses in word examples were put not on vowels, but on syllables, as it was in the Cyrillic, and previously described Latin transcriptions. Putting such an accent complies with the international format.

Table 4
Transformation of words into the Latin transcription (according to the IPA) (an excerpt)

Allophone	Word example (1)	Latin transcription symbol (IPA)	Word example (2)
C'0	пацо+кваць.	t□□	p□□□□kv□t□□
CH0	трагі+чны.	□	tr□□□□i□n□

Thus, we have developed the knowledge base for the system of transcription generation for the Belarusian language. The knowledge base makes possible to put a symbol of one of three desirable phonetic alphabets in correspondence with any allophone which can be generated by the system of text-to-speech synthesis.

Conclusion

This paper says about the beginning of work on the development of a system for the producing of different transcriptions for any input electronic texts in the allophonic representation in the Belarusian language. Its experimental prototype has been implemented as a free service, constantly available on the Internet at <http://www.corpus.by/convertAllophToCyrPhonemes/> together with the resource www.corpus.by [4, 10]. With its help, one can automatically receive real-time transcription of any Belarusian allophonic text pieces according to three phonetic alphabets.

The service quickly transforms a text according to the Cyrillic or Latin transcription, understandable to anyone since school. Converting into the international IPA transcription opens up the possibility for foreigners to facilitate understanding of the phonetics basics of a Belarusian written text. In addition to the direct use of the service as a tool for learning the Belarusian language, it becomes also possible to quickly and automatically create specialized Belarusian phonetic dictionaries which can significantly help phoneticians. Besides, using this service together with the text-to-speech synthesizer allows a user to control and correct the work of the synthesizer with the help of notifications sent by e-mail. Further improvement of the system will lie in the integration of the system with the output of the allophonic processor of the text-to-speech synthesizer, and identification of the errors found. It is also planned to modify the algorithm of placing the correct accents (on a stressed syllable, not on a stressed vowel) for the IPA transcription.

We would like to thank Alena Skopinava for her help in revising the language of this paper.

References

1. Text-to-Speech PHP-Based Synthesizer [Electronic resource]. – 2013. – Mode of access : <http://corpus.by/tts3>. – Date of access : 08.12.2013.
2. Open Semantic Technology for Intelligent Systems [Electronic resource]. – 2010. – Mode of access : <http://www.ostis.net>. – Date of access : 08.12.2013.

3. Чахоўскі, Г.К. Сучасная беларуская мова. Фанетыка. Фаналогія. Арфаэпія / Г.К. Чахоўскі, Т.Л. Чахоўская. – Мінск : БДУ, 2010. – 110 с.
4. Гачко, А.К. Метадычныя рэкамендацыі па раздзелу “Фанетыка. Фаналогія” курса “Сучасная беларуская мова”: фанетычная транскрыпцыя [Электронны рэсурс] / А.К. Гачко. – 2012. – Рэжым доступу : http://edu.grsu.by/books/gachko_phonetic/index.php/fanetychnaya-transkryptsyua. – Дата доступу : 09.12.2013.
5. Беларуская-англійскі размоўнік / уклад. У. А. Кошчанка. – Мінск : Артэя груп, 2010. – 190 с.
6. The International Phonetic Association [Electronic resource]. – 2005. – Mode of access : <http://www.langsci.ucl.ac.uk/ipa/index.html>. – Date of access : 08.12.2013.
7. Wiktionary: International Phonetic Alphabet [Electronic resource]. – 2008. – Mode of access : http://en.wiktionary.org/wiki/Wiktionary:International_Phonetic_Alphabet. – Date of access : 08.12.2013.
8. Гецэвіч, Ю.С. Распрацоўка сінтэзатара беларускага і рускага маўленняў па тэксце для мабільных і інтэрнэт-платформаў / Ю.С. Гецэвіч, Д.А. Пакладок, Д.В. Брэк // Развитие информатизации и государственной системы научно-технической информации (РИНТИ-2012) : доклады XI Междунар. конф., Минск, 15 ноября 2012 г. – Минск : ОИПИ НАН Беларуси, 2012. – С. 254–259.
9. Лобанов, Б.М. Компьютерный синтез и клонирование речи / Б.М. Лобанов, Л.И. Цирульник. – Минск : Белорусская наука, 2008. – 342 с.
10. Service of converting allophonic texts into different transcriptions = Сэрвіс канвертавання алафоннага тэксту ў розныя транскрыпцыі [Electronic resource]. – 2013. – Mode of access : <http://corpus.by/convertAllophToCyrPhonemes>. – Date of access : 08.12.2013.

DISTINGUISHING MENTAL ACTIVITY TYPES BY ELECTROENCEPHALOGRAPHY RECORDS

D. Kanonik

United Institute of Informatics Problems of the NAS of Belarus, Minsk

e-mail: denisbsu@gmail.com

In this paper a technique is presented to classify electroencephalography (EEG) records, taken from subjects during periods of different kinds of mental activity. It was possible to compare described approach with neural networks classifier for AR coefficients presented in [1]. Fitting a general model for all patients leads to a high error rate, but an individual approach allows building high quality classifiers for each of the patients. It means that it's possible to build a functional system that can figure out what kind of mental activity a subject performs. Such system will be useful in any kind of critical industry or military control process, where it's vital that the operator is awake and involved. Also it may help person to maintain concentration on subject.

Introduction

Current state of technology enables researchers to record low noise EEG with highly portable and convenient devices. As EEG represents state of brain electrical activity that is highly correlated with person's state of mind, this device may be used to determine the kind of mental activity a person is doing. As that person doesn't need to have any special training, brain-computer interface (BCI) devices based on mind state are much more convenient to use. The low number of distinguishable states and slow (for inexperienced operator) transition between states does not allow building a high precision and fast BCI system, but for a number of applications such approach may be useful. The only downside of this method of interaction with a computer is that the person should actually perform some sort of mental activity to transmit the command, but, for example, in entertainment applications this is completely acceptable. Also, while monitoring a subject's state (e.g. determining whether operator of dangerous machinery is actually awake and involved in the process), this property of system is even desirable.

There are numerous special properties of an EEG signal, both time and frequency domain-related. In the frequency domain historically the first to be observed was the division of the spectrum into four frequency bands (delta, theta, alpha and beta), each one related to some type of neuron activity. The downside of this technique is its high sensitivity to noise. In the time domain the autoregressive model [2] was found to be very useful for separating normal subjects from those with psychiatric disorders [3]. In [1], the autoregression (AR) model coefficients were fed to a neural network and yielded rather good results. The idea of this work is to improve on their results even further by merging the AR coefficients and frequency bands power values to increase the accuracy of the method. Spatial separation was natural - since an electrode is more sensitive to nearer events, features for each electrode may be calculated separately and then merged together.

Some form of machine learning should be used to analyze those features. In [1] neural networks showed good results, and also provided some insight about inner cerebral processes. In this research an SVM-based classifier shows good results as a replacement of neural networks.

There was a high risk of overfitting the models because of huge number of features, so some form of dimensionality reduction required. Best results were obtained while using independent components analysis (ICA) to perform this reduction.

Data Source

Data was obtained previously by Keirn and Aunon [4, 5]. The subjects of the experiment were seated in an Industrial Acoustics Company sound controlled booth with dim lighting and noiseless fans for ventilation. Signal from electrodes C3, C4, P3, P4, O1 (20-10 electrode placement system) of an Electro-Cap electrode cap was passed through Grass 7P511 amplifiers with a 0.1-100 Hz bandpass filter. The signal was digitized by a Lab Master 12 bit analog-to-digital converter at 250Hz sampling rate.

The analog filter has a sharp enough cutoff below the Nyquist frequency, and the frequency bands necessary for this research were far below it, no additional digital filtering was needed.

For this paper, data from seven subjects performing five mental tasks was analyzed. Those five tasks are:

- the baseline task (subjects were asked to relax);
- the letter task (subjects were asked to compose a letter to friend or relative without vocalizing it);

- the math task (subjects were asked to solve nontrivial multiplication problems);
- the visual counting task (subjects were asked to imagine a blackboard and to visualize numbers being written on it);
- the geometry rotation task (subjects were asked to visualize a particular 3D block figure, being rotated around an axis).

Data was recorded for 10 seconds during each task. The tasks were repeated 5 times per session. Two sessions for each person was taken for analysis.

Obtaining Features

Each record was divided into segments 1 second long that overlap by 0.5 seconds. Then segments containing eye blinks were discarded. Blinks were detected using a record from a special pair of eye electrodes - sharp slopes on this signal correspond to blinks. Next step is re-referencing signal - to get rid of electrical drift and other induced signals, new reference signal was calculated as the average signal of all electrodes and all data was recalculated according this new reference. After this step of data rectification, data features were computed. For every segment of the record 3 AR coefficients and 7 band powers (1-4 Hz, 4-7 Hz, 7-9 Hz, 9-10.5Hz, 10.5-13 Hz, 13-20 Hz, 20-30 Hz, normalized to have total sum 1) are computed for each channel.

Autoregression coefficients were calculated using the R ARIMA model implementation from the stats library with no difference term and no intercept term, so the fitted model had the following form:

$$X[i] = a_1 \cdot X[i-1] + a_2 \cdot X[i-2] + a_3 \cdot X[i-3], \quad (1)$$

where $X[i]$ is the value of the rectified signal at the i -th sample, a_i are the autoregression coefficients.

Band coefficients were obtained using the R implementation of fft from the fftw library, then summing up all frequencies magnitudes in each bracket and dividing each sum by the overall sum for all brackets.

As data provides 5 different channels, this method yields 50 features for each segment. Using these features directly in the machine learning stage leads to poor performance, as with such a number of features and a relatively small dataset, machine learning has strong tendencies to overfitting. To move ahead, dimensionality of the feature space should be reduced.

Dimensionality Reduction

After a set of experiments with various dimensionality reduction techniques, ICA has been selected. Intuitively, as ICA could transform data in a way that lines up most “nongaussian” dimensions, it should provide nearly the best separable space within chosen number of dimensions. In a series of experiments this intuition was confirmed, as best results was achieved with ICA reduction to 11 dimensions for SVM and 15 for neural networks. Reduction was performed by the fastICA R implementation, using the training set as input and producing translation from the 50D feature space to the 11D (15D) target space.

Classification

Two approaches to classification were tested: neural networks and support vector machine (SVM). For a representation of neural networks, the multilayer perceptron was chosen. With two hidden layers and 15 units in each layer this perceptron provides fast computation speed while maintaining worthy classification results. The RSNNS library from R provided the implementation for it. Competing with this neural network was an SVM implemented in the kernlab R library with the Crammer, Singer native multi-class classifier.

To determine the performance of each classifier cross-validation was used to limit the overfitting problem. Also the low amount of data suggests using cross-validation to get rid of “lucky” and “unlucky” cases. To cover this part of implementation, simple k-fold cross-validation was implemented (25-fold specifically)

In each round of cross-validation both classifiers was trained on the test set, consisting of records tagged with the class of mental activity and divided into segments. In the classification stage for every record it was split in segments, then each segment was classified by the classifier to produce a “vote” with

relevant class. Then the overall record class was determined by the simple majority of votes. This way the classifier, trained on segments, was able to classify entire records. This approach was inspired by [1] and it reduced error rate even further. Performance of the selected classifiers highlighted in table.

Classification quality after cross-validation

	Subject 1	Subject 2	Subject 3	Subject 4	Subject 5	Subject 6	Subject 7
NN	0.82	0.8	0.46	0.7143	0.4533	0.8	0.64
SVM	0.78	0.84	0.52	0.6735	0.4667	0.68	0.68

Conclusion

Results in this paper significantly improve on results from [1]. It means that additional frequency features can improve accuracy of BCI system significantly. Also as both a small perceptron network and an SVM produced rather similar results, both methods may be used, depending of implementation availability in a particular BCI device. High recognition quality without any feedback or subject training may be useful in various fields.

References

1. Anderson, C.W. Classification of EEG Signals from Four Subjects During Five Mental Tasks / C.W. Anderson, Z. Sijercic. – Department of Computer Science, Colorado State University, 1996.
2. Evaluation of parametric methods in EEG signal analysis / S-Y. Tseng [et al.] // Med. Eng. Phys. – 1995. – Vol. 17. – P. 71–78.
3. Tsoi, A.C. Classification of EEG using artificial neural networks / A.C. Tsoi, D.S.C. So, A. Sergejew // Advances in Neural Information Processing Systems 6. – 1994. – P. 1151–1158.
4. Keirn, A.Z. Alternative modes of communication between man and machine / A.Z. Keirn // Master's thesis. – Purdue University, 1988.
5. Keirn, A.Z. A new mode of communication between man and his surroundings / A.Z. Keirn, J.I. Aunon // IEEE Transactions on Biomedical Engineering. – 1990. – Vol. 37 (12) – P. 1209–1214.

PLAIN OBJECTS DETECTION IN IMAGE BASED ON A CONTOUR TRACING ALGORITHM IN BINARY IMAGES

A. Kazlouski, R.K. Sadykhov
Belarusian State University of Informatics and Radioelectronics, Minsk
e-mail: anton.kazlouski@yahoo.com

Objects detection in image is a central problem to many tasks in digital image processing and therefore it has a vast range of applications. An algorithm for plain objects detection in image based on a contour tracing algorithm in a binary image is presented in this paper. The detection of plain object in image relates to the decomposition of a binary image contour into contours and closed contours and is based on the developed by the author algorithms for: an applicant for the vertex detection in a binary image, a closed contour detection in a binary image. The algorithm for an applicant for the vertex detection in a binary image allows performing the decomposition of a binary image contour into contours and closed contours. The problem of the determination of the starting point for contour tracing algorithm in a binary image is considered. The concept of plain object in image allows analyzing images of different types with a large number of considered classes of plain objects. The experiments show the robustness and the efficiency of the algorithm.

Introduction

There are many object detection techniques in various image processing applications such as remote sensing, technical vision, medicine, and many others. It is one of the fundamental and still open image processing tasks due to the complexity of analysis of images of different types with a large number of considered classes of objects. It relates to image interpretation problem.

Current object detection techniques [1–4] can be categorized into *bottom-up*, *top-down*, and combining the previous two approaches relates to the template matching approach. *Bottom-up* approaches [1, 4] start from low-level or mid-level image features, i.e. edges (contours, boundaries) or segments. These methods build up hypotheses from such features, extend them by construction rules and then evaluate by certain cost functions [3].

Top-down approaches [2] use object representation learned from examples based on prior knowledge about an object, such as shape, color, or texture, to guide the segmentation. Their result is approximate due to complexity of accurate description of classes of objects.

The third category of approaches combining *top-down* and *bottom-up* methods [3] have become prevalent because they take advantage of both aspects. The result is as close as possible to the top-down approximation, but is also constrained by the bottom-up process to be consistent with significant image discontinuities.

The problems of contour detection and segmentation are related. In general, contour detectors offer no guarantee that they will produce closed contours and hence do not necessarily provide a partition of the image into regions. Closed contours can be recovered from regions in the form of their boundaries. The result of image segmentation is a set of segments that cover the input image [5, 6]. Another way of closed contours detection in image is the decomposition of a binary image contour into contours and closed contours. It refers to starting and stopping criteria of a contour tracing algorithm in a binary image. The current approach is to start trace a binary image contour from the first encountered pixel of contour. It does not allow detecting a closed contour without additional complex analysis. As the result, the problem of an arbitrary contour decomposition is still opening.

A new algorithm for plain objects detection in image was developed and evaluated by the author. The detection of plain object in image refers to the developed by the author algorithms for: an applicant for the vertex detection in a binary image, a closed contour detection in a binary image. The algorithm for an applicant for the vertex detection in a binary image allows performing the decomposition of a binary image contour into contours and closed contours. The concept of plain object in image allows analyzing images of different types with a large number of considered classes of plain objects.

1. Plain object in image

A *contour* in a digital image is a connected set of pixels determined from the given input image according to a predefined set of rules.

Edges must be linked into a representation for region boundary and contour. Contour can be open or closed. A *closed contour* correspond to a region boundary, as a boundary of a finite region in image forms a closed path.

A *plain object* in a digital image is an object, whose decomposition on n closed binary image contours is not possible, where $n \in \mathbf{N}$ and $n > 1$. The boundary of the plain object does not belong to a set of decomposition of k closed binary image contours of another object, where $k \in \mathbf{N}$ and $k > 1$. Herewith, the boundary of the plain object cannot be represented as a decomposition of sets of l binary image contours and p closed binary image contours, where $p, l \in \mathbf{N}$ and $p, l \geq 1$. The binary image is derived from the given input image according to a predefined set of rules.

The conditions of the definition of a plain object in image are agreed according to the possibility to link contours in a binary image into a segment, and sometimes the segment is linked in such a way that correspond to a boundary, but this is not always the Plain objects are illustrated in (fig. 1).

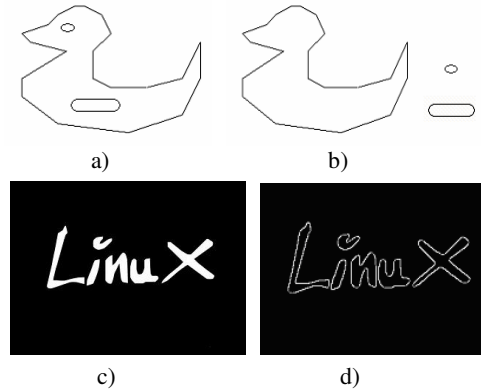


Fig. 1. Images: a) an object of the image whose decomposition on n closed contours possible, $n = 3$; b) a set of closed contours extracted from the image, see fig. 1, a; c) an image; d) a binary image derived from the image, see fig. 1, c (all the objects are plain)

2. An algorithm for plain objects detection in image based on a contour tracing algorithm in a binary image

There are three main steps of the considered algorithm for plain objects detection in image: 1) an image binarization; 2) closed contours detection in a binary image; 3) a plain object identification.

Image binarization is based on the Canny edge detection algorithm [7]. The result is a binary image with “thin edges”, must be 1-pixel contour.

The developed by the author algorithm for a closed contour detection in a binary image scans the given input binary image line-by-line from left to right, and top to bottom. Herewith, the starting point for a contour tracing algorithm is determined based on the developed by the author algorithm for an applicant for the vertex detection. A closed contour is determined by the developed stopping criteria of a binary image contour tracing algorithm.

The determination of a binary image pixel as an applicant for the vertex relates to the corresponding algorithm whose scheme illustrated in fig. 2. The pixels marked with black squares in fig. 2 will call pixels of forming angle.

A binary image pixel $f(j, i)$ is an applicant for the vertex of the first type in case of equality of pixels of forming angle $f(j, i+1)$ and $f(j+1, i)$ to one (marked with black squares, fig. 2, a). Herewith, the value of pixel $f(j-1, i+2)$ or $f(j+2, i-1)$ is equal to zero (marked with 0, fig. 2, a) and pixels values, marked with 2 or 3, do not equal to one simultaneously. There are no pixels of forming angle from left and right or top and bottom of the pixels of forming angle [8].

A binary image pixel $f(j, i)$ is an applicant for the vertex of the second type in case of equality of pixels of forming angle $f(j, i-1)$ and $f(j+1, i)$ to one (marked with black squares, fig. 2, b). Herewith, the value of pixel $f(j-1, i-2)$ or $f(j+2, i+1)$ is equal to zero (marked with 0, fig. 2, a) and pixels values, marked with 2 or 3, do not equal to one simultaneously. There are no pixels of forming angle from left and right or top and bottom of the pixels of forming angle [8].

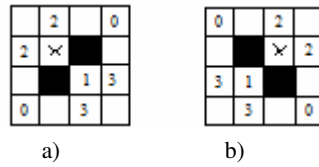


Fig. 2. The schemes of a binary image pixel determination as an applicant for the vertex: a) first type; b) second type

Plain object identification is based on the analysis of the localization results of the closed contour detection algorithm. The localization of a closed contour relates to the determination of the extreme left, right, top, bottom contour pixels when tracing the contour. Herewith, a considered closed contour must not have more than n_3 spur contours with the length more than n_4 pixels. It is achieved based on the algorithm for an applicant for the vertex detection in a binary image and a contour tracing algorithm in a binary image.

The algorithm parameters are as follows: n_1 – the maximum number of binary image pixels to be traced; n_2 – the minimum length (in binary image pixels) of a closed binary image contour, which represents a boundary of a plain object; n_3 – the maximum number of spur binary image contours of a plain object boundary; n_4 – the maximum number of binary image pixels to be traced within the detection of spur binary image contours.

3. An algorithm for a closed contour detection in a binary image

There are three starting rules of a binary image contour tracing algorithm: an applicant for the vertex must be determined to start the tracing procedure; the binary image contour tracing procedure starts from the angle forming pixel $f(j+1, i)$; a binary image pixels with respect to the applicant for the vertex ($f(j, i)$): $f(j+1, i+1)$ or $f(j+1, i-1)$ do not take part in the tracing procedure, i.e. must be equal to 0 (marked with 1, fig. 2).

There are three stopping rules of a binary image contour tracing algorithm: the binary image contour tracing procedure stops if the end of a contour has been reached, i.e. there is no possibility for further tracing. Note, the contour tracing algorithm must not allow to trace contour vice versa, i.e. the angle forming pixel $f(j+1, i)$ must not meet during tracing. The binary image contour is not closed; the binary image contour tracing procedure stops if n_2 image pixels have been successfully traced. The binary image contour is not closed; the binary image contour tracing procedure stops if within it one of the following pixels have been met with respect to the applicant for the vertex ($f(j, i)$): $f(j, i-1)$, $f(j, i)$, $f(j, i+1)$. The binary image contour is closed (see fig. 2).

The proposed algorithm for a closed contour detection in a binary image is independent from a contour tracing algorithm (might be used anyone).

4. Experimental results

The considered algorithm for plain objects detection in image was tested with the following parameters: $n_1 = 1200$, $n_2 = 50$, $n_3 = 2$, $n_4 = 15$. The function *bwboundaries*, MATLAB [9] was tested with the following parameters values: *conn* = 8, *options* = 'noholes'. The function *bwconncomp*, MATLAB [10] was tested with a default connectivity of 8. The extraction of contours relates to the Canny edge detection algorithm. The Canny edge detection algorithm was used with the following thresholds values: $T_1 = 0$ and $T_2 = 0,3$. The k-mean clustering segmentation was tested with the following parameters values: *distance* = 'sqEuclidean', $k = 3$.

The tests results are estimated based on the following coefficients: $k_1 = N/R$ и $k_2 = L/R$, where N – the number of undetected plain objects in the image; L – the number of false plain objects detected in the image; R – the total number of plain objects in the image, $R = 1, \dots, m, m \in \mathbf{N}$.

Fig. 3 shows a test sample T of 5 images of different types. Fig. 4 shows K-mean clustering segmentation results. The test results are presented in table 1, 2, 3.

Table 1

The test results of functions *bwboundaries*, matlab / *bwconncomp*, MATLAB

I	Detected contours	Detected plain objects	Undetected plain objects	False plain objects	Total number of plain objects	Detection time, c	kk_1	kk_2
1	2	3	4	5	6	7	8	9
I_1	123	0	18	123	18	0,0544/0,007	1	6,8

I_2	21	0	2	21	2	0,0223/0,002	1	10,5
I_3	6	6	0	0	6	0,0089/0,011	0	0
I_4	23	0	5	23	5	0,0070/0,002	1	4,6
I_5	24	0	7	24	7	0,0270/0,002	1	3,4

Table 2

The test results of the considered algorithm for plain object detection, MATLAB

I	Detected closed contours	Detected plain objects	Undetected plain objects	False plain objects	Total number of plain objects	Detection time, c	kk_1	kk_2
1	2	3	4	5	6	7	8	9
I_1	23	18	0	0	18	1,7328	0	0
I_2	1	1	2	1	2	0,5041	1	0,5
I_3	6	6	0	0	6	0,0390	0	0
I_4	5	2	3	1	5	0,7520	0,6	0,2
I_5	10	8	0	3	7	0,0773	0	0,4

Table 3

The test results of k-mean clustering segmentation, MATLAB

I	Detected segments	Detected plain objects	Undetected plain objects	False plain objects	Total number of plain objects	Detection time, c	kk_1	kk_2
1	2	3	4	5	6	7	8	9
I_1	27	8	12	0	18	7,6	0,7	0
I_2	10	3	0	1	2	3,2	0	0,5
I_3	6	0	6	0	6	3,3	1	0
I_4	14	7	0	2	5	2,8	0	0,4
I_5	0	0	7	0	7	3,4	1	0



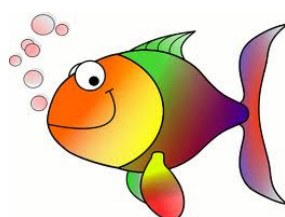
a)



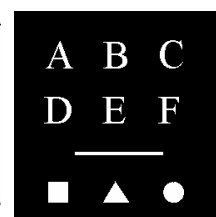
b)



c)



d)



e)

Fig. 3. The test sample T of 5 images of different types: a) an image I_1 , size of 501×501 pixels; b) an image I_2 , size of 256×256 pixels; c) an image I_3 , size of 320×240 pixels; d) an image I_4 , size of 259×194 pixels; e) an image I_5 , size of 252×252 pixels

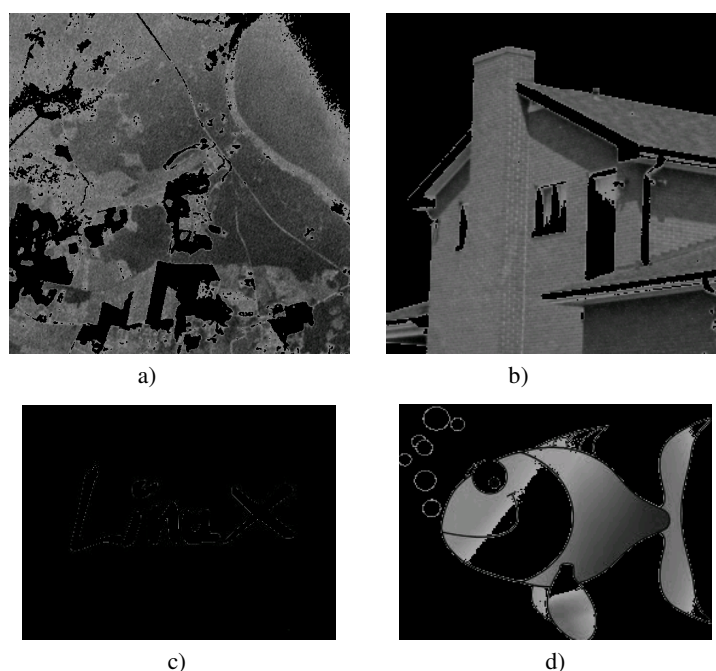


Fig. 4. K-mean clustering segmentation results, see fig. 3

Conclusion

The new algorithm for plain objects detection in image based on a contour tracing algorithm in a binary image is introduced. The use of the algorithm for an applicant for the vertex detection in a binary image allows performing decomposition of a binary image contour into contours and closed contours without a significant increase in processing time. A transition to processing of plain objects in image allows analyzing images of different types with a large number of considered classes of plain objects. The results of experiments confirmed the efficiency of the algorithm.

References

1. Recovering human body configurations: Combining segmentation and recognition / G. Mori [et al.] // Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. – Washington, DC, USA, 2004. – Vol. 2. – P. 326–333.
2. Borenstein, E. Learning to Segment / E. Borenstein, S. Ullman // Proceeding of the 8th European Conference on Computer Vision. – Prague, Czech Republic, 2004. – Vol. 3. – P. 315–328.
3. Object Detection Combining Recognition and Segmentation / L. Wang [et al.] // Proceedings of the 8th Asian Conference on Computer Vision. – Tokyo, Japan, 2007. – Vol. 1. – P. 189–199.
4. Contour analysis introduction and its image and signal processing application / Ya.A. Furman [et al.]. – M. : Nauka, 2002. – 592 p. [in Russian]
5. Contour Detection and Hierarchical Image Segmentation / P. Arbel'aez [et al.] // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2011. – Vol. 33, № 5. – P. 898–916.
6. Luo, M. Ma. A spatial constrained K-means approach to image segmentation / M. Ma Luo, H.Y. Zhang // Proceedings of the 4th Pacific Rim Conference on Multimedia. – 2003. – Vol. 2. – P. 738–742.
7. Canny, J. Computational Approach to Edge Detection / J. Canny // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 1986. – Vol. 8, no. 6. – P. 679–698.
8. Kazlouski, A. A corner detector based on an approximation of contours of an image's objects / A. Kazlouski // Informatics. – 2010. – № 4. – P. 36–47.
9. Mathworks [Electronic resource]. – Mode of access : <http://www.mathworks.com/help/images/ref/bwboundaries.html>. – Date of access : 16.04.2014.
10. Mathworks [Electronic resource]. – Mode of access : <http://www.mathworks.com/help/images/ref/bwconncomp.html>. – Date of access : 16.04.2014.

DEVELOPMENT OF DOMAIN-SPECIFIC LANGUAGE FOR CONSTRUCTING DYNAMIC SUBJECT DOMAINS

H. Karkanitsa

Yanka Kupala State University of Grodno, Belarus

e-mail: karkanica@gmail.com

The article continues a series of papers devoted to the technology of development a dynamic subject domain based on distributed expert knowledge. The present paper describes the use of Domain-Specific Modeling (DSM) methodology to automate the process of constructing subject domain in decision support systems. A concept of development of domain-specific language is presented.

Introduction

Efficiency and quality of administrative solutions directly depend on possession of innovative knowledge. In this connection tasks of acquisition, representation and operative usage of innovative knowledge become especially actual. Solution of these tasks complicates the fact that in the modern informational environment knowledge, especially technological, possesses property of distribution and has a short life cycle.

In decision support systems knowledge is formalized according to some model, is stored in knowledge bases and is applied in decision-making processes. Efficiency of decision-making is reached by means of engaging of experts who can in the shortest periods of time provide relevancy, reliability and entirety of knowledge. Experts, as a rule, are dispersed worldwide. Their knowledge actually is distributed cognitive resource. So we are faced with the problem of constructing a subject domain which is, first, is formed from a variety of distributed resources, and secondly, is dynamic. It is necessary to develop automation technology of constructing subject domain for such class of problems.

Earlier, dynamic subject domain model based on the graph theory was constructed as the theoretical basis of the technology [1, 2]. The present paper considers the development of domain-specific language (DSL) which is designed to express statements in a particular problem space, or domain.

1. Problem statement

Let us formulate the formal description of the task on technology development. Let T - the task demanding an operational decision. We will define it, as a collection of following components: $T = (S, Group, S_1, \dots, S_n, Req_1, \dots, Req_k, Solution_1, \dots, Solution_n, t_1, \dots, t_n)$, where S - task setting; $Group = (Center, Expert_1, Expert_2, \dots, Expert_n)$ - the group of performers including center ($Center$), initiating the task, and the distributed performers ($Expert$), implementing solution of subtasks; S_1, \dots, S_n - settings of the subtasks received as a result of decomposition S ; Req_1, \dots, Req_k - requirements to solution; $Solution_1, \dots, Solution_n$ - the formalized information received as a result of solution of subtasks; t_1, \dots, t_n - limitations of span time on solution of subtasks; $n \rightarrow \infty, k$ - a variable value.

As a basis for construction of model of a subject domain of task T the graph theory is used, allowing simulating hierarchy of subtasks. As model of a subject domain of task T we will understand attribute tree $G(V, E)$. Within the limits of the given model expert knowledge is presented in attributes of vertices.

According to model of task T to every tree vertex the following set of attributes is put in correspondence: $v = \langle id, task, name, addr, status, inf \rangle$, where id - the unique identifier of vertex (task); $task$ - task setting (requirements to solution); $status$ - state of vertex (the task is initiated, the task is in process of solution, the task is solved); $name$ - the unique identifier of the expert; $addr$ - the address of the expert; inf - informational component (actually the solution of the task presented in one of admissible formats). It is necessary to develop domain-specific language (DSL) specialized to a particular subject domain of task T that must have a reasonably small set of elements that can be easily defined and extended to represent domain-specific constructs.

2. Development of DSL, the base concepts

There are several disadvantages of universal modeling languages such as UML. UML is designed to support mostly object oriented programming. Consequently, in contrast to domain-specific, UML is used for a wide variety of purposes across a broad range of domains. It has primitives whose semantics

are not familiar to all experts in particular domain. To use universal modeling language experts need to know their syntax and semantic.

Identified disadvantages of universal languages led to a new modeling approach (Domain Specific Modeling) which involves the development of specific, domain-specific modeling languages created for the particular subject domain. Unlike general-purpose languages, DSL is more expressive, easy to use and understandable to experts.

In our case it is required to develop the language of representation of the dynamic graph model of task *T*. It gives us opportunity to describe the model as a formal specification that permits machine representation of the graph.

A domain-specific language must combine the set of elements and the relationships between them into a coherent grammar. It must also define whether combinations of elements and relationships are valid. Fundamental concepts of graph domain model of task *T* are nodes (Node) and relationships (Relationship) between them. Nodes have a numeric identifier (index).

Relationships organize vertices into a hierarchical tree structure and combine set of nodes-subtasks (Nodes). Relationship has an initial node (Start Node), the end node (End Node) and type. Each relationship has a type (RelationshipType) which is uniquely identified by a string name (name). Nodes and relationships have properties (Properties). Properties are used to set the task's attributes. Properties are defined as a "key-value" pair (fig. 1).

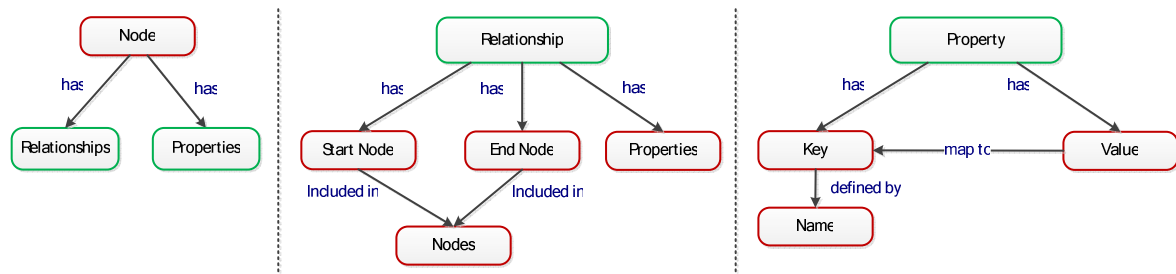


Fig. 1. DSL element's graphical notation

Graph consists of nodes and relationships combined their associated properties in which data are stored. The node in this case is a representation of the vertices in the graph (task, subtask). Link-element defines a hierarchy of subtasks (fig. 2).

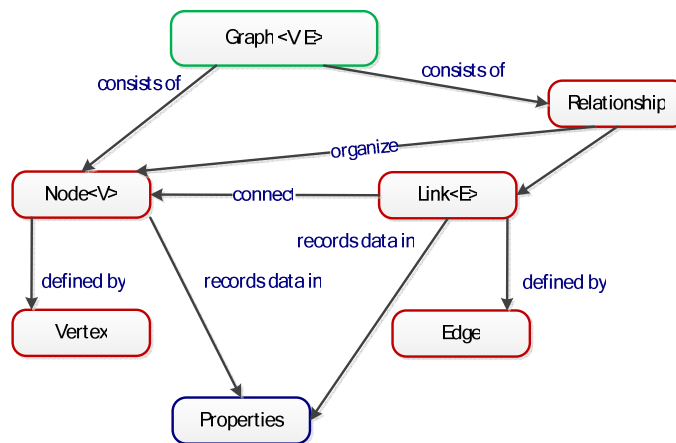


Fig. 2. Graphical notation of Graph element

3. DSL for dynamic subject domain. Language statements

Let's consider a group of statements that must be included in the domain-specific language. First, according to the dynamic property of the domain model we should include statements that can be used to perform basic operations (create node or relationships, add node or relationships, delete node or group of nodes and others): create-operators, delete-operators, return-operators. They are used to implement algorithms of defragmentation and synthesis of the domain model. Secondly, it is necessary to define a group of statements that allow setting and changing the attributes of nodes and relationships (set-operators). Thirdly, the language must contain a set of statements common to all high-level programming languages: conditional operator, loop operator and others. This group of operators will be used to implement algorithms for processing and modification of the domain model.

DSL statements syntax are based on understandable interpretation of concepts established in practice and on a simple notation which is borrowed from the concept of SQL, as well as the approach to the implementation of the query language of graph-oriented database Neo4j. Some statements are represented (table).

Table

Statement	Action	Example
create n={ ... }	creates a single node and set an attributes	create n = {id : '1.1', task : 'task1.1', name:'expert1', status : 0}
create a-[r:relype]->b	creates a relationship between two nodes	start a=node(1), b=node(2) create a-[r: includes]->b
create ({list})	creates a collection of nodes	create {1.1, 1.2, 1.3}
start n=node(<index> return n	returns the node identified by index	start n=node (1) return n
start r=relationship(<id> return r	returns a relationship with the id	start r=relationship(0) return r
start n=node({list}) return n	returns a collection of nodes	start n=node(1, 2, 3) return n
start n=node(*) return n	returns all nodes	start n=node(*) return n
delete	delete a single node/ relationship/ attribute	start n = node(4) delete n
set	set an attribute on a node or relationship	set n.name = 'Expert1'
foreach	allows to do updating commands on elements in a collection	start begin = node(2), end = node(1) foreach(n in nodes(p) : SET n.status = 1)

In terms of defined set of statements we can construct the subject domain model and implement algorithms of its processing and modification.

Conclusion

The main result is following. The domain-specific language for representing the graph model of the dynamic subject domain is developed. As a result, the description of the scene of problem solution is significantly simplified. Algorithms are formulated in terms of a particular domain without using unnecessary notations and terminologies of universal modeling languages. A limited set of elements and operators simplifies implementation of the required algorithms of constructing and modifying the model of dynamic subject domain.

References

1. Karkanitca, A. Development of Dynamic Subject Domain Based on Distributed Expert Knowledge / A. Karkanitca // Modeling and Simulation : MS'2012 : Proc. of the Intern. Conference. – Minsk : Publ. Center of BSU, 2012. – P. 137–139

2. Karkanitsa, H. Combined Method of Knowledge Acquisition for Constructing DMSS / H. Karkanitsa, N. Dzeyeva // AWERProcedia Information Technology & Computer Science. – 2013. – Vol. 3. – P. 559–564.

STATISTICAL ANALYSIS OF SPATIO-TEMPORAL DATA BY POISSON CONDITIONAL AUTOREGRESSIVE MODEL

Yu. Kharin, M. Zhurak
Research Institute for Applied Problems of Mathematics and Informatics,
Belarusian State University, Minsk
e-mail: kharin@bsu.by

A new model for spatio-temporal data is proposed. Statistical estimators for the model parameters are constructed. Statistical forecasting algorithms to predict the states for future time moments are proposed and their risks are evaluated. The theoretical results are illustrated in computer experiments.

Introduction

The models based on spatio-temporal statistical data are widely used in statistical analysis of real processes. They are used in meteorology, ecology, economics, medicine and other fields [1]. Statistical analysis and modeling spatio-temporal data is a challenging task. There is a need to process a large amount of data to build adequate models. Analyzing the spatio-temporal data, often it is impossible to find a closed form, and it requires using big computer resources. Some models of spatio-temporal observations are natural generalization of the univariate time series or spatial data models. However, in order to construct models that adequately describe the real data new approaches are often required for modeling the spatial and temporal effects simultaneously.

1. Poisson conditional autoregressive model

Let's introduce the notation: (Ω, F, P) is the probability space; N is set of positive integers, $N_0 = N \cup \{0\}$; $S = \{1, 2, \dots, n\}$ is a set of spatial objects (sites), n is a number of spatial objects; $t \in \mathfrak{S} = \{1, 2, \dots, T\}$ is discrete time, T is the length of observation period; $x_{s,t} \in N_0$ is a discrete random variable at time t at site s (e.g., the incidence rate); $U(s) \subset \{1, 2, \dots, s-1\} \subseteq S$ is a subset of sites which most closely situated to the site s ($U(1) = \emptyset$); $F_{\bar{s}, < t} = \sigma\{x_{u,\tau} : u \neq s, \tau < t\} \subset F$ is the σ -algebra generated by the indicated in braces random variables; $z_{s,t} \geq 0$ is an observed (known) level of exogenous factors (e.g., environmental pollution) at time t at site s which influences $x_{s,t}$; $\{\varphi_k(t) : 1 \leq k \leq K\}$ is a given set of $K \in N$ basic functions which determine a trend; $L\{\xi\}$ means the probability distribution of the random variable ξ ; $E\{\xi\}$ is the expectation of the random variable ξ ; $Po(l; \lambda)$ is the Poisson probability distribution with the parameter $\lambda > 0$:

$$P\{\xi = l\} = Po(l; \lambda) = \frac{\lambda^l}{l!} e^{-\lambda}, \quad l \in N_0. \quad (1)$$

We construct the Poisson conditional autoregressive model for spatio-temporal data $\{x_{s,t}\}$ [2]:

$$L\{x_{s,t} | F_{\bar{s}, < t}\} = Po(x_{s,t}; \lambda_{s,t}), \quad (2)$$

$$\ln \lambda_{s,1} = \ln \lambda_{s,1} \left(\{x_{j,1} : j \in U(s)\} \right) = \sum_{j \in U(s)} b_{sj} x_{j,1} + \beta_s z_{s,1} + \sum_{k=1}^K \gamma_{sk} \varphi_k(1), \quad s \in S, \quad (3)$$

$$\ln \lambda_{s,t} = \ln \lambda_{s,t} \left(\{x_{j,t} : j \in U(s)\}, x_{s,t-1} \right) = a_s x_{s,t-1} + \sum_{j \in U(s)} b_{sj} x_{j,t} + \beta_s z_{s,t} + \sum_{k=1}^K \gamma_{sk} \varphi_k(t), \quad (4)$$

$t = 2, 3, \dots, T, s \in S,$

where $a = (a_1, a_2, \mathbf{K}, a_n)' \in R^n$, $b_s = (b_{s_1}, \mathbf{K}, b_{s_{|U(s)|}})' \in R^{|U(s)|}$, $j_k \in U(s) : k = 1, \dots, |U(s)|$, $s \in S$, $\beta = (\beta_1, \beta_2, \mathbf{K}, \beta_n)' \in R^n$, $\gamma_s = (\gamma_{s_1}, \mathbf{K}, \gamma_{s_K})' \in R^K$, $s \in S$ are the parameters of the model. The number of parameters of the model is equal to $D = n(2 + K) + \sum_{s=1}^n |U(s)|$.

Let $X_t = (x_{1,t}, x_{2,t}, \mathbf{K}, x_{n,t})' \in N_0^n$ be a vector of values of observed variables for all sites at time t . The model (1)-(4) represents the Markov chain model of the dependence on time and on components of the vector X_t . Let $L = \{l_j = (l_{j,1}, \mathbf{K}, l_{j,n})' \in N_0^n : j = 1, 2, \mathbf{K}\}$ be an ordered set of all admissible values of the vector X_t . Following the model (1)-(4) X_t is the nonhomogeneous Markov chain with the one-step transition probability matrix:

$$P(t) = (p_{ij}(t)),$$

$$p_{ij}(t) = P\{X_t = l_j | X_{t-1} = l_i\} = Po(l_{1,j}; \lambda_{1,t}(l_{1,i})) \prod_{s=2}^n Po(l_{s,j}; \lambda_{s,t}(\{l_{k,j} : k \in U(s)\}, l_{s,i})), \quad t \geq 2.$$

An initial probability distribution $p = (p_0, p_1, \mathbf{K})'$ of the Markov chain X_t by (1)-(4) is defined as follows:

$$p_j = P\{X_1 = l_j\} = Po(l_{1,j}; \lambda_{1,1}) \prod_{s=2}^n Po(l_{s,j}; \lambda_{s,1}(\{l_{k,j} : k \in U(s)\})).$$

As the conditional probability distribution we use the Poisson distribution (1) with intensity $\lambda_{s,t}$ which is defined by (3), (4) and depends on $\{x_{j,t} : j \in U(s)\}$ and $x_{s,t-1}$ if $t = 2, 3, \mathbf{K}, T$ and on $\{x_{j,1} : j \in U(s)\}$ if $t = 1$.

By the Kolmogorov-Chapman formula the matrix of transition probabilities $H(t_1, t_2) = (h_{ij}(t_1, t_2))$, $h_{ij}(t_1, t_2) = P\{X_{t_2} = l_j | X_{t_1} = l_i\}$ for $t_2 - t_1$ steps from time point t_1 to time point t_2 ($t_1 < t_2$) has the form:

$$H(t_1, t_2) = P(t_1 + 1)P(t_1 + 2)\mathbf{K} P(t_2). \quad (5)$$

The current probability distribution $p(t) = (p_0(t), p_1(t), \mathbf{K})'$, $p_j(t) = P\{X_t = l_j\}$, $j = 1, 2, \mathbf{K}$, at time t can be find using the total probability formula:

$$p_j(t) = \sum_{i=0}^{\infty} p_i h_{ij}(1, t), \quad l_j \in L, t \in N. \quad (6)$$

2. Statistical estimation of parameters

Introduce the notation: $\theta_s = (a_s, b_s', \beta_s, \gamma_s')' \in R^{2+K+|U(s)|}$, $\theta = (\theta_1, \mathbf{K}, \theta_n)' \in R^{n(2+K)+\sum_{s=1}^n |U(s)|}$ is a composite vector of unknown parameters which needs to be estimated.

Theorem 1. The log-likelihood function for the model (1)-(4) for an observed set of spatio-temporal data $\{X_t : t = 1, 2, \dots, T\}$ takes the additive form:

$$l(\theta) = \sum_{s=1}^n l_s(\theta_s), \quad (7)$$

$$l_s(\theta_s) = \sum_{t=1}^T (-\lambda_{s,t} + x_{s,t} \ln \lambda_{s,t} - \ln x_{s,t}!), \quad (8)$$

where $\lambda_{s,t}$ is defined by (3), (4).

To find the maximum likelihood estimators (MLE) of the parameters of the model we need to maximize the log-likelihood function (7), (8):

$$l(\theta) \rightarrow \max. \quad (9)$$

Based on the model (1)-(4) and Theorem 1 the problem (9) splits into n maximization problems:

$$l_s(\theta_s) \rightarrow \max_{\theta_s}, s \in S. \quad (10)$$

A necessary condition for a local maximum in (10) is

$$\nabla_{\theta_s} l_s(\theta_s) = \mathbf{O}_{2+K+|U(s)|}, \quad (11)$$

where $\mathbf{O}_{2+K+|U(s)|}$ is a vector $2+K+|U(s)|$ zero elements.

We solve (11) numerically using the Newton iterative method which has the quadratic convergence. For this method the $(k+1)$ th iteration has the form ($k = 0, 1, 2, \dots, K$):

$$\theta_s^{(k+1)} = \theta_s^{(k)} - \left(\nabla_{\theta_s}^2 l_s(\theta_s^{(k)}) \right)^{-1} \cdot \nabla_{\theta_s} l_s(\theta_s^{(k)}),$$

where $\theta_s^{(k)}$ is an approximation of the MLE $\hat{\theta}_s$ on the k -th step, $\nabla_{\theta_s} l_s(\theta_s^{(k)})$ is the vector of the first order derivatives at the point $\theta_s^{(k)}$, $\nabla_{\theta_s}^2 l_s(\theta_s^{(k)})$ is the matrix of the second order derivatives at the point $\theta_s^{(k)}$. The iterative computation stops if $\|\nabla_{\theta_s} l_s(\theta_s^{(k)})\| < \varepsilon$, where $\varepsilon \geq 0$ is a given small quantity which determines the computation accuracy; in this case we take the statistic $\hat{\theta}_s = \theta_s^{(k+1)}$ as a solution.

3. Statistical forecasting

Consider the problem of forecasting of the future state $X_{T+\tau}$ in $\tau \geq 1$ steps ahead based on observations until the time $t=T$ inclusively: X_1, \mathbf{K}, X_T . Denote forecasting statistic $\hat{X}_{T+\tau} = g_\tau(X_1, \mathbf{K}, X_T; \theta)$, where θ is the vector of true values of the model (1)-(4) parameters. We characterize the error of forecasting of the future state $X_{T+\tau}$ in $\tau \geq 1$ steps ahead based on the statistic $g_\tau(X_1, \mathbf{K}, X_T; \theta)$ using the matrix mean square risk of forecasting [3]:

$$r(\tau) = E \left\{ \left(\hat{X}_{T+\tau} - X_{T+\tau} \right) \left(\hat{X}_{T+\tau} - X_{T+\tau} \right)' \right\} \in R^{n \times n}.$$

Define the conditional matrix mean square risk for the forecast $\hat{X}_{T+\tau}$ provided that the entire prehistory $\{X_t : t=1, \dots, T\}$ before the moment T inclusively is known:

$$r_c(\tau, X_1, X_2, \mathbf{K}, X_T) ::= E \left\{ \left(\hat{X}_{T+\tau} - X_{T+\tau} \right) \left(\hat{X}_{T+\tau} - X_{T+\tau} \right)' \mid X_1, X_2, \mathbf{K}, X_T \right\} \in R^{n \times n}. \quad (12)$$

Since the fact that in the model (1)-(4) time dependence is defined only on 1 step the conditional risk (12) depends only on X_T :

$$r_c(\tau, X_1, X_2, \mathbf{K}, X_T) = r_c(\tau, X_T) = E \left\{ \left(\hat{X}_{T+\tau} - X_{T+\tau} \right) \left(\hat{X}_{T+\tau} - X_{T+\tau} \right)' \mid X_T \right\}.$$

Then the unconditional risk can be calculated as follows:

$$r(\tau) ::= E \{ r_c(\tau, X_T) \} = \sum_{k=0}^{\infty} r_c(\tau, l_k) P \{ X_T = l_k \}. \quad (13)$$

Using the formula (6) of the probability distribution at time T the formula (13) transforms to the form:

$$r(\tau) = \sum_{k=0}^{\infty} \sum_{i=0}^{\infty} r_c(\tau, l_k) p_i h_{ik}(1, T).$$

We construct the optimal forecast $\hat{X}_{T+\tau}^{(1)}$ minimizing the mean square risk, the iterative forecast $\hat{X}_{T+\tau}^{(2)}$ and the forecast $\hat{X}_{T+\tau}^{(3)}$ minimizing the probability of the forecast error in case when parameters θ of the model (1)-(4) are known.

Theorem 2. Optimal (in terms of minimum of the mean square risk) forecast by T previous observations $\{X_t : t=1, K, T\}$ for the model (1)-(4) in $\tau \geq 1$ steps ahead has the form

$$\hat{X}_{T+\tau}^{(1)} = \sum_{j=0}^{\infty} l_j h_{ij}(T, T+\tau), \quad (14)$$

where $l_j \in L (j=1, 2, K)$, $h_{ij}(T, T+\tau)$ can be found using (5), i is determined by the equation $l_i = X_T$.

Corollary. Conditional mean square risk for the forecasting statistic (14) provided that the entire prehistory $\{X_t, t=1, K, T\}$ is given has the following form:

$$r_{c1}(\tau, X_T) = \sum_{j_1=0}^{\infty} \sum_{j_2=0}^{\infty} l_{j_1} (l_{j_1} - l_{j_2})' h_{ij_1}(T, T+\tau) h_{ij_2}(T, T+\tau).$$

Since $L\{x_{s,t} | F_{\bar{s}, < t}\} = Po(x_{s,t}; \lambda_{s,t})$, then $E\{\hat{x}_{s,t} | F_{\bar{s}, < t}\} = \lambda_{s,t}$. Considering this property of the Poisson distribution the forecast in $\tau \geq 1$ steps ahead for the model (1)-(4) can be calculated iteratively:

$$\hat{x}_{s, T+h}^{(2)} = \exp \left\{ a_s \hat{x}_{j, T+h-1}^{(2)} + \sum_{j \in U(s)} b_{sj} \hat{x}_{j, T+h}^{(2)} + \beta_s z_{s, T+h} + \sum_{k=1}^K \gamma_{sk} \Phi_k(T+h) \right\}, \quad h=1, K, \tau, \hat{x}_{s, T}^{(2)} = x_{s, T}, s \in S.$$

Lemma. Conditional mean square risk for the forecast $\hat{X}_{T+\tau}^{(2)}$ provided that the entire prehistory $\{X_t, t=1, K, T\}$ is given has the following form:

$$r_{c2}(\tau, X_T) = \sum_{j=0}^{\infty} \left(\hat{X}_{T+\tau}^{(2)} \hat{X}_{T+\tau}^{(2)'} + l_j l_j' - \hat{X}_{T+\tau}^{(2)} l_j' - l_j \hat{X}_{T+\tau}^{(2)'} \right) h_{ij}(T, T+\tau), \quad l_j \in L, j=1, 2, K,$$

where $\hat{X}_{T+\tau}^{(2)} = (\hat{x}_{1, T+\tau}^{(2)}, \hat{x}_{2, T+\tau}^{(2)}, \dots, \hat{x}_{n, T+\tau}^{(2)})' \in R^n$, $h_{ij}(T, T+\tau)$ can be found using (5), i is determined by the equation $l_i = X_T$.

Since X_t is the Markov chain, as in [3], as a measure of forecast accuracy we can use also the probability of the forecast error:

$$r(\tau) = P\{\hat{X}_{T+\tau} \neq X_{T+\tau}\}.$$

Theorem 3. The optimal forecasting statistic minimizing the probability of the forecast error $r(\tau)$ is as follows:

$$\hat{X}_{T+\tau}^{(3)} = l_{\arg \max_j \{h_{ij}(T, T+\tau)\}}, \quad (15)$$

where i is determined by the equation $l_i = X_T$. In this case the minimum of the risk is

$$r(\tau) = 1 - \sum_{j=0}^{\infty} \max_i \{h_{ij}(T, T+\tau)\} P\{X_T = l_i\} = 1 - \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \max_i \{h_{ij}(T, T+\tau)\} p_k h_{ki}(1, t).$$

Note that if there is more than one maximum element in (15), then we have the equivalent forecasts which have the same risk.

In case of parametric prior uncertainty we construct forecasting statistics using “the plug-in principle” [3]:

$$\hat{X}_{T+\tau}^{\mathcal{C}} = g_{\tau}(X_1, K, X_T; \hat{\theta}),$$

where $\hat{\theta}$ is the MLE of the parameters of the model (1)-(4) based on observations $\{X_1, K, X_T\}$.

4. Results of computer experiments

Computer experiments are performed on simulated data. We consider the model (1)-(4) with the following values of parameters: $n=3$, $K=2$, $S = \{1, 2, 3\}$, $U(1) = \emptyset$, $U(2) = \{1\}$, $U(3) = \{2\}$, $\varphi_1(t)=1$, $\varphi_2(t)=t$, $\beta_s \equiv 0$, $\theta_1 = (-0.1, 1.6, -0.005)'$, $\theta_2 = (-0.2, -0.07, 1.7, -0.01)'$, $\theta_3 = (-0.1, -0.08, 2.1, -0.01)'$, $T = 20$. Statistical estimators of the parameters obtained by $T = 20$ observations have the form:

$$\hat{\theta}_1 = (-0.12, 1.95, -0.046)', \quad \hat{\theta}_2 = (-0.35, -0.03, 1.29, -0.02)', \quad \hat{\theta}_3 = (-0.08, -0.24, 2.24, -0.001)'$$

Fig. 1 plots the dependence of the mean square error $\hat{v} = \hat{E}\{\|\hat{\theta} - \theta\|^2\}$ estimated by $M = 1000$ Monte Carlo realizations, on the length T of the observation period ($T \in [10, 220]$).

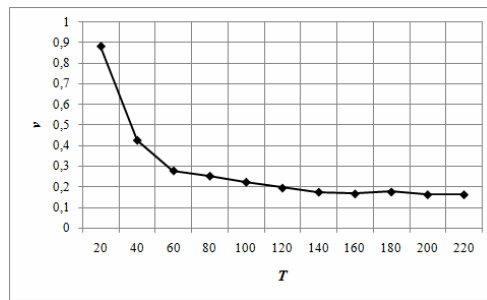


Fig. 1. Dependence of the mean square risk on T

Fig. 2 shows simulated data and computed “plug-in” forecasts in $\tau = 5$ steps ahead at future time points $t \in \{21, 22, 23, 24, 25\}$ for $s = 1$.

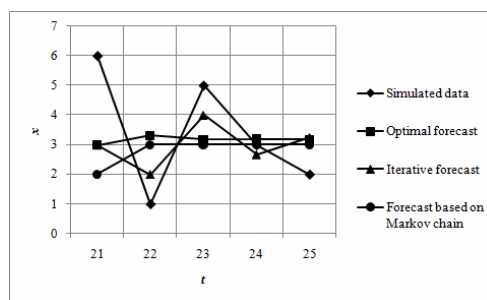


Fig. 2. Forecasting results for the site $s = 1$

Conclusion

In this paper the Poisson conditional autoregressive model for spatio-temporal data is proposed. An algorithm to compute the maximum likelihood estimators for the model parameters is developed. Optimal forecasting statistic minimizing the mean square risk, the iterative forecasting statistic and the forecasting statistic minimizing the probability of the forecast error are constructed and their risks are evaluated. The results of computer experiments on simulated data are presented.

References

1. Handbook of Spatial Statistics / edited by Alan E. Gelfand [et al.]. – Taylor and Francis Group, LLC, 2010 – 608 p.
2. Kharin, Yu.S. Poisson’s conditional autoregressive model and its estimation based on spatio-temporal data / Yu.S. Kharin, M.K. Zhurak // Proceedings of NAS of Belarus. Physics and Mathematics series. – 2013. – Vol. 3. – P. 22–30. [In Russian]
3. Kharin Yu. Robustness in Statistical Forecasting / Yu. Kharin. – N.Y. : Springer, 2013. – 356 p.

HIERARCHICAL PIXEL CLUSTERING FOR IMAGE SEGMENTATION

M. Kharinov

St. Petersburg Institute for Informatics and Automation of RAS

e-mail: khar@iias.spb.su

In the paper a piecewise constant image approximations of sequential number of pixel clusters or segments are treated. A majorizing of optimal approximation sequence by hierarchical sequence of image approximations is studied. Transition from pixel clustering to image segmentation by reducing of segment numbers in clusters is provided. Algorithms are proved by elementary formulas.

Introduction

The paper focuses on the domain of image segmentation by optimal approximations that minimally differ from the image of N pixels in the standard deviation σ or total squared error $E = N\sigma^2$. Although related approaches, namely Otsu methods [1, 2], K -means method [3] and Mumford-Shah model [4–7] have a long history, the opportunities of minimizing of the total squared error E are far from being exhausted, especially, in the task of multiple optimization for each number of pixel clusters or, in particular, connected image segments. In this task Otsu's multi-thresholding [2] provides an accurate but incomplete solution for clustering of pixels. Mumford-Shah model [4–7] provides a complete sequence of image partitions into each number of segments, but minimizing effect is poor. K -means method for image segmentation is too heuristic to provide any of mentioned two requirements, but it can be advanced for application in conjunction with Otsu method and Mumford-Shah model [8].

To solve the task of multiple optimization without any difficulties we use a special data structure of Sleator-Tarjan dynamic trees [9] that essentially optimizes the computing, but does not affect the obvious meaning of algorithms. Therefore, to avoid the cumbersome details of implementation, here we address rather to motivation of solutions and do not dwell on the software that supports the fast generation, storing in the available RAM and effective transformations of pixel clusters in a computer memory.

To substantiate the study of segmentation results without appealing to the subsequent detection of a priori specified objects, we have calculated the optimal and nearly optimal approximations for the simplest examples of real images [10]. These proved important for the formulation of the problem caused by two challenges.

1. Problem statement

The first challenge is that a sequence of optimal image approximations in general case is not hierarchical [10]. But just hierarchical sequence of approximations is quite accessible for computational optimization. Therefore, there arises the problem of majorizing of none-hierarchical optimal approximation sequence by quasioptimal hierarchical sequence of approximations, which don't significantly differ from the optimal ones in total squared error E or standard deviation σ .

The second challenge is illustrated by two-level approximations in fig. 1.

Leftmost in fig.1 is the original image, central is the optimal approximation with two intensities obtained by conventional Otsu method, providing the minimum of standard deviation of approximation from the image. Rightmost is the nearly optimal approximation of original image by two segments. The values of standard deviation are written under the approximations.

Comparing the two approximations in fig.1, it is easy to notice that in nearly optimal approximation the segments of optimal approximation are connected to each other by natural or artificial coupling elements one pixel wide. The calculations of such coupling elements are unstable. With increasing intensity resolution the contribution of coupling elements in the total squared error tends to zero. Then, none-literal “virtual” coupling elements turn out to be preferable, and we come to utilization of pixel clusters instead of less general connected segments.



Fig. 1. Two-level optimal and nearly optimal approximations of the standard Lenna image

In this case, the problem is to obtain the complete sequence of the quasioptimal image approximations by pixel clusters with a limited number of segments. The decisive argument to treat clusters of connected and disconnected pixels instead of less common connected segments consists in that for clusters, besides a sole merging operation, two additional operations, which don't cause withdrawal from their sets, are introduced.

2. Elementary formulas for total squared error minimizing

To optimize image approximations by the total squared error E or standard deviation σ , we use three operations with pixel clusters, namely, merging, splitting and correction, described by the following formulas.

Let I_1 and I_2 be the average intensities for clusters 1 and 2, respectively. Let n_1 be the number of pixels in the cluster 1 and n_2 be the number of pixels in the cluster 2. Then the increment ΔE_{merge} of the total squared error E caused by the merging of specified clusters along with reduction of the number of clusters per unit is given by the formula:

$$\Delta E_{merge} = \frac{(I_1 - I_2)^2}{\frac{1}{n_1} + \frac{1}{n_2}} \geq 0. \quad (1)$$

Just the increment ΔE_{merge} is minimized in the version [4] of Mumford-Shah model [5], however, in application to connected segments. In the version [6], the appropriate formula differs by an additive term, and in FLSA version [7] by a multiplicative factor to take into account the total length of the boundaries between the segments (clusters of connected pixels).

Let's write down the formula for splitting of the cluster 1, when its $k < n_1$ pixels with average intensity I initiate a new cluster. In this case, the cluster 1 is split into two clusters of k and complementary $n_1 - k$ pixels, and cluster splitting is accompanied with increase of the cluster number per unit along with a non-positive increment ΔE_{split} of the total squared error:

$$\Delta E_{split} = -\frac{(I - I_1)^2}{\frac{1}{k} - \frac{1}{n_1}} \leq 0. \quad (2)$$

For cluster splitting it is important that in (2) a predetermined set of pixel clusters is expected, which obtained, say, by cluster merging using (1). So, the extended set of clusters is taken into account when splitting. Another feature of algorithms based on the formula (2), is the update of the hierarchy of clusters, which is performed for each of the nested clusters, treated both as the individual images.

The composition of splitting and merging of clusters induces a correction operation without changing the number of clusters, which is accompanied with an increment $\Delta E_{correct}$ of the total squared error:

$$\Delta E_{correct} = \frac{(I - I_2)^2}{\frac{1}{k} + \frac{1}{n_2}} - \frac{(I - I_1)^2}{\frac{1}{k} - \frac{1}{n_1}}, \quad (3)$$

where the negative term in (3) describes the increment of the total squared error E , caused by converting of k pixels from cluster 1 into a separate cluster, and the first term in (3) describes the increment of E caused by merging of the initiated cluster with the cluster 2, in accordance with (1) and (2).

A notable feature of (3) is that by its simplifying the method K -means is derived [8]. Applying (3) precisely, we have proposed for the clustering of pixel sets a more accurate method [10], which provides a calculation of a complete sequence of optimal image approximations that are treated in multi-threshold Otsu method [2]. Another feature of formula interpretation in this paper is the application of (3) for generation of a hierarchy of pixel clusters, unlike overlapping partitioning, as in [8, 10].

3. Quasioptimal image approximations

To produce the sequences of quasioptimal approximations, we follow the principle of dichotomous division of the cluster into two subclusters independently of the others, treating each cluster as a separate image. In this way the splitting of non-uniform clusters containing the different pixels is performed, while clusters of all identical pixels are treated as indivisible or elementary. To avoid analysis of cluster repetitions the computation of hierarchical sequence of quasioptimal image approximations, containing 1, 2, 3, ... clusters of pixels, is performed in two stages. At the first stage so called "compact" invariant representation, which specifies the sequence of partitions of the image pixels into 1, 2, 4, 8 ... clusters, is calculated. At the second stage, a compact representation is expanded into a sequence of approximations with successively increasing numbers of clusters, so as to provide the maximal decrease of the total squared error E or standard deviation σ .

In an obvious way the sequence of quasioptimal image approximations is obtained by splitting of the non-uniform clusters according to conventional histogram Otsu method [2], wherein the threshold intensity value is found using exhaustive search, from the condition of maximum decrease of the total squared error E .

In a more complicated algorithm, the quasioptimal approximations are generated according to the formula (2), using a hierarchy of clusters that previously generated in bottom-up strategy by brute-force implementation of formula (1) for the minimizing of total squared error E over all pairs of clusters. As we have established experimentally, the results of calculations according (1) and (2) coincide with each other. Moreover, the equivalent hierarchy of pixel clusters can be generated in the simplest algorithm [11], if we exchange the heuristic criterion [11] of merging of successive histogram bins by criterion $\Delta E_{merge} = \min$, where ΔE_{merge} is detailed in (1).

Characteristic feature of optimal and quasioptimal image approximations with $g = 1, 2, 3, 4, \dots$ number of pixel clusters is that the corresponding sequence $E_1 \geq E_2 \geq \dots \geq E_g$ of values of the total squared error E is convex:

$$E_i \leq \frac{E_{i-1} + E_{i+1}}{2}, \quad i = 2, 3, \dots, g - 1. \quad (4)$$

Convexity property (4) holds also for a compact sequence of quasioptimal image approximations with $g = 1, 2, 4, 8, \dots$ clusters. Thus, the quasioptimal approximations preserve a convexity property of the optimal image approximations.

Fig. 2 shows the image approximations with two, three and four pixel clusters, visualized by the same number of average intensities. Comparing fig. 2 with fig. 1, it is easy to notice the visual difference between two types of quasioptimal approximations that is less expressed in E or σ values.

4. Correcting of image approximations by segment number

The main limitation of quasioptimal approximations fig. 2 in image segmentation task is a sharp increase in the number of connected segments with increasing number of pixel clusters. However, this

limitation is overcome due to the reducing of the number of segments in the image approximation that is performed as an additional step in algorithm of generation of quasioptimal approximations by cluster splitting into two subclusters.



Fig. 2. Quasioptimal image approximations with 2, 3, 4 pixel clusters

Reducing the number of segments is performed by reclassifying pixels of non-isolated segment from donor subcluster to the acceptor subcluster that are selected to minimize the increment of the total squared error $\Delta E_{correct} = \min$, where $\Delta E_{correct}$ is detailed in (3).

Fig. 3 illustrates the quality σ of approximations of the standard image, depending on the number of clusters g shown in the range from one to one thousand.

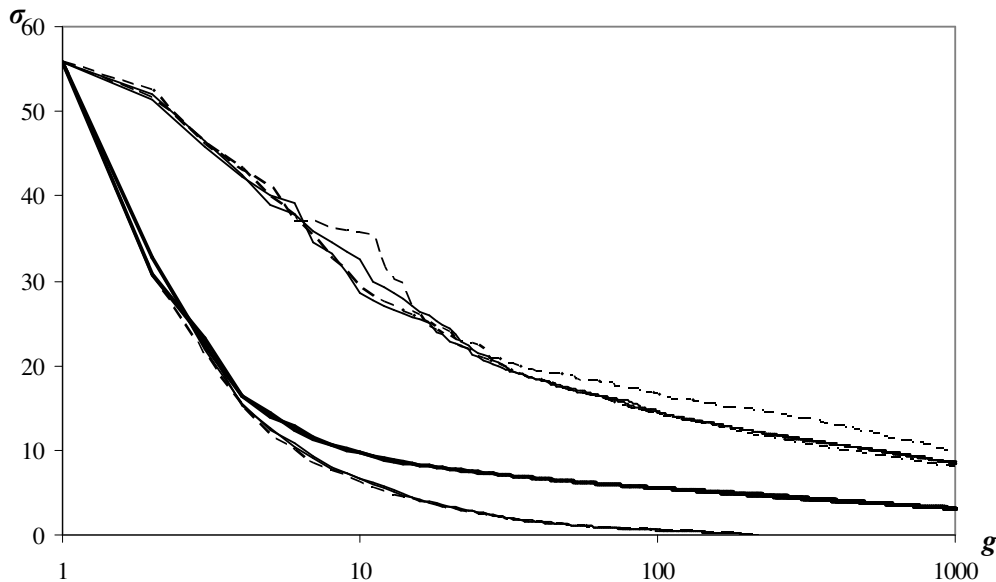


Fig. 3. Standard deviation σ depending on the number of clusters g (logarithmic scale)

In fig. 3 the pairs of intertwining solid curves describe the sequence of approximations obtained by top-down clustering according Otsu method and bottom-up clustering by iterative merging. Lower boundary dotted curve marks the optimal approximations with successive number of pixel clusters. The pair of almost merged solid curves just above this dashed curve describes two majorizing sequences of quasioptimal image approximations. Upper dashed curve describes image approximations with successive number of connected segments according to FLSA version [7] of Mumford-Shah model, and just below dashed curve corresponds to the version [4] of Mumford-Shah model [5].

The uppermost pair of intertwined solid curves describes the sequence of approximations obtained in discussed bottom-up and top-down algorithms, along with the reduction of the number of segments in each subcluster to one. In this case the sequence of image approximation with sequential number of segments is generated, as in the Mumford-Shah model. Intermediate pair of bold curves describes a sequence of approximations obtained by reduction of the number of segments, which terminates under certain stopping condition. In discussed particular case a uniqueness of average intensity of each segment was treated as a stopping condition that causes three-five times reduction in the segment number, compared to quasioptimal approximations. For a variety of stopping conditions the area between the curves in fig. 3 becomes available that extends the capabilities of image segmentation via pixel clustering.

Conclusion

Thus, the quasioptimal image approximations fig. 2 may be reproduced by the histogram algorithm [11], wherein the «distance between the clusters», i.e. the product of intra-class and inter-class variances, is to be replaced by ΔE_{merge} from (1)². As we have established experimentally, this algorithm provides the minimization, which remains valid for the general case of merging of any cluster pairs. Along the way, we have created software for the joint analysis of global and local pixel features, which came in handy to develop a clustering method by reduction of the number of segments.

It should be noted that in addition to the availability of computing, the quasioptimal image approximations have one more remarkable advantage, compared to optimal approximations. Concretely, the quasioptimal approximations are easily converted into invariant image representations [12] that don't depend on the linear transformations of pixel intensities. This topic should be discussed in the following papers.

References

1. Otsu, N. A Threshold Selection Method from Gray-Level Histograms / N. Otsu // IEEE Transactions on systems, MAN, and CYBERNETICS. – 1979. – Vol. SMC-9 (1). – P. 62–66.
2. Liao, Ping-Sung A Fast Algorithm for Multilevel Thresholding / Ping-Sung Liao, Tse-Sheng Chen, Pau-Choo Chung // J. Inf. Sci. Eng. – 2001. – Vol. 17(5). – P. 713–727.
3. Jain, A.K. Data Clustering: 50 Years Beyond K -Means / A.K. Jain // Pattern Recognition Letters. – 2010. – Vol. 31 (8). – P. 651–666.
4. Bugaev, A.S. Pilot studies and development of methods and tools for analysis and automatic recognition of media flows in global information systems / A.S. Bugaev, A.V. Khelvas // Scientific and Technical Report. – Vol. 1. – Moscow : MIPT, 2001.– 140 p. [in Russian].
5. Mumford, D. Boundary detection by minimizing functionals, I / D. Mumford, J. Shah // Proc. IEEE Comput. Vision Patt. Recogn. Conf. – San Francisco, 1985. – P. 22–26.
6. Koepfler, G. A Multiscale Algorithm for Image Segmentation by Variational Method / G. Koepfler, C. Lopez, J. Morel // SIAM J. on Numerical Analysis. – 1994. – Vol. 31 (1). – P. 282–299.
7. Robinson, B.J. Implementation of a fast algorithm for segmenting SAR imagery / B.J. Robinson, N.J. Redding, D.J. Crisp // Scientific and Technical Report. – Australia : Defense Science and Technology Organization, 2001. – 42 p.
8. Kharinov, M.V. A generalization of three approaches to an optimal segmentation of digital image / M.V. Kharinov // SPIIRAS Proceedings. – 2013. – Vol. 2 (25). – P. 294–316.
9. Galiano, P. Application of Sleator-Tarjan Dynamic Trees in a Monitoring System for the Arctic Region Based on Remote Sensing Data / P. Galiano, M. Kharinov, S. Vanurin // Information Fusion and Geographic Information Systems (IF AND GIS 2013). – Springer, 2014. – P. 137–147.
10. Kharinov, M.V. Image segmentation method by merging and correction of sets of pixels / M.V. Kharinov // Pattern Recogn. Image Anal.: Adv. Math. Theory Appl. – 2013. – Vol. 23 (3). – P. 393–401.
11. Arifin, A.Z. Image segmentation by histogram thresholding using hierarchical cluster analysis / A.Z. Arifin, A. Asano // Pattern Recogn. Letters. – 2006. – Vol. 27 (13). – P. 1515–1521.
12. Kharinov, M.V. Storage and adaptive processing of digital image information / M.V. Kharinov. – St. Petersburg University Press, 2006. – 138 p. [in Russian].

² The optimal approximations of the standard Lenna image, which may be desired for comparison, are available at <http://oogis.ru/content/view/107/42/lang.ru/>

IMAGE ANALYSIS IN HIGH-VELOCITY INTERACTION OF A GROUP OF BODIES WITH A TARGET

I.E. Khorev^{1,2}, S.A. Zelepugin^{1,3}, V.F. Tolkachev¹, A.S. Zelepugin¹

¹ Tomsk State University, Russia;

² Tomsk State University of Control Systems and Radioelectronics, Russia;

³ Tomsk Scientific Centre, Siberian Branch RAS

The processes of deformation and destruction of finite-thickness targets during the high-velocity interaction with a group of compact impactors were experimentally investigated and numerically simulated. Three-dimensional computations with the use of the finite-element method were performed within the framework of an elastoplastic model of the medium. The final stages of the formation of through holes were simulated using complete destruction criteria. Simple criteria for estimating the mutual influence of the impactors and the character of destruction of the target at various incidence angles and distances between the impactors were obtained. Peculiarities of image analysis in the problem of high-velocity interaction were mentioned.

Introduction

The mathematical simulation of physical interaction between targets and a group of variously sized particles imitating a stream of technogenic fragments in near space is a promising method of gaining reliable data about the dynamics of the total process of a group impact and its consequences [1-5]. The comprehensive experimental and theoretical analysis of this complicated problem must involve the following components: the development of methods and devices for the launching of a group of particles under laboratory conditions in air and in vacuum, the experimental investigation of an impact of particles with targets and its consequences for protected samples, and numerical simulation with development of an adequate closed procedure for computation of the impact of a group of fragments with targets imitating the protection of space equipment and directly with the space-apparatus construction.

1. Launching of a Group of Particles

Among a large variety of possible systems of controlled launching of a fragment stream under laboratory conditions, assemblies using the aerodynamic principle of the step-by-step separation of a launched construction are of considerable interest, because they require no additional power supplies to provide a given orientation of fragments in the group. In this case, it is necessary to organize the process of separation of various trays and leading facilities in a possibly short time interval.

In this study, we use the launching of a group of particles (from 2 to 12) in air on the basis of the separation of compound systems which were composed of identical bodies sequentially so that their longitudinal axes coincided or were parallel to each other and to the longitudinal axis of the whole system. In this case, we provided the process of the directed ejection of fragments from a container under the action of aerodynamic forces.

Fig. 1, 2 show shadow photographs for the motion of four and six spherical bodies, respectively, at the Mach number $M = 3.3$. The spatial positions of the particles corresponded to their primary arrangement in the container, where four particles were disposed in the vertices of a regular quadrangle and six particles were arranged as two identical triples at the container ends, i.e., formed a “dumb-bell”.

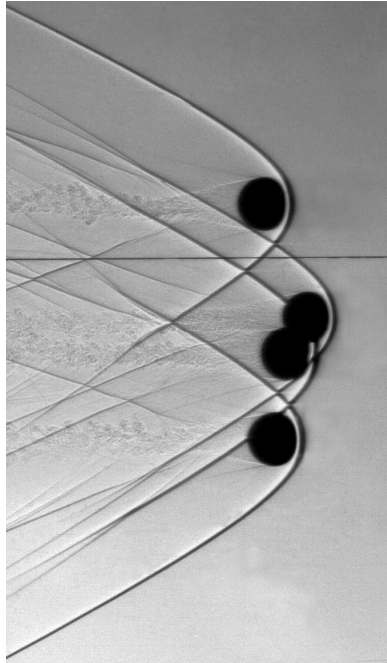


Fig. 1. Shadow photograph of the motion of a groups of four spherical particles

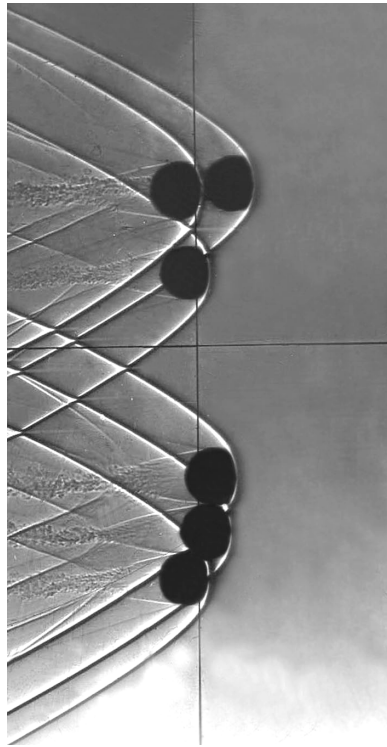


Fig. 2. Shadow photograph of the motion of a groups of six spherical particles

2. Experimental Investigation

The experimental investigations of the processes of interaction in the groups of bodies were carried out in a light-gas ballistic installation with subsequent analysis of the preserved specimens. The first series of the tests was aimed at the evaluation of the spall area in obstacles of finite thickness depending on the density of the uniform flow of high-speed impactors. As impactors, we used steel balls. The total number of balls in the flow varied from 2 to 7. We used four types of balls were with diameters of 3, 4, 6, and 8 mm and masses of 0.11, 0.26, 0.88, and 2.1 g, respectively. The initial velocity of the impactors was chosen close to 3000 m/s. At these velocities, all materials undergo fracture and we observe the formation of shock waves and multiple spall phenomena. In all tests, the role of targets was played by steel plates 6, 8, and 10 mm in thickness.

In fig. 3, we present the experimentally obtained views of the back side of an 8-mm-thick obstacles after the frontal collision of three and four impactors 6 mm in diameter at an initial velocity of 2873 m/s. The effect of the mutual influence of the processes of penetration of each impactors on the fracture of the obstacle strongly depends on the initial distances between the strikers. In the last case, we observe the formation of a region of back-side spall common for all four impactors. The first three through holes in the obstacle merge into a single hole but the last through hole formed by the right impactor remains isolated.

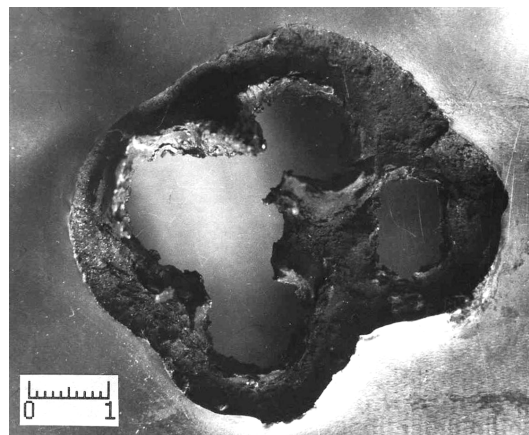
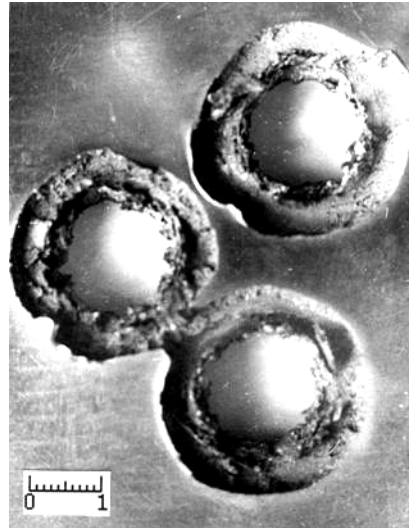


Fig. 3. Back-side views of the obstacles after the impact of groups of three and four elements

In the second series of the tests, we studied the variations of the total mass of back-side spall in the obstacle for a constant total mass of homogeneous impactors in the flow set equal to 20 g. As a reference case, we used the case of a single impactor with a mass of 20 g. The thickness of the obstacle was equal to 10 mm and the initial impact velocity was equal to 3000 m/s. The mass of spall (the ratio of the total mass of the spall fragments to the initial mass of the impactors) significantly increases with the number of impactors in the flow. This effect is close to the effect observed in the case of transition from a striker of compact shape to a striker in the form of a plate of equivalent mass. Since, in the case of a plate, the initial area of contact with the obstacle is larger, the greater fraction of the kinetic energy of the plate is spent for the spall fracture of the obstacle, rather than for the formation of a through hole in the target and the residual velocity of flow behind the obstacle. In the case where the role of spall fracture is predominant, a simple procedure of fragmentation of the initial high-speed impactor results in a strong intensification of the effect of spall fracture in the obstacle.

3. Numerical Simulation

To numerically simulate high-velocity loading, we used a damageable-medium model that postulates the existence of microcavities (pores and cracks). The total volume of the medium was

assumed to be composed of the compact component and microcavities (the density of which is set equal to zero). The system of equations describing the nonstationary adiabatic movement of the medium (for both elastic and plastic deformation) with consideration given to the evolution of microdefects was composed of the equations of conservation of mass, momentum, and energy [6].

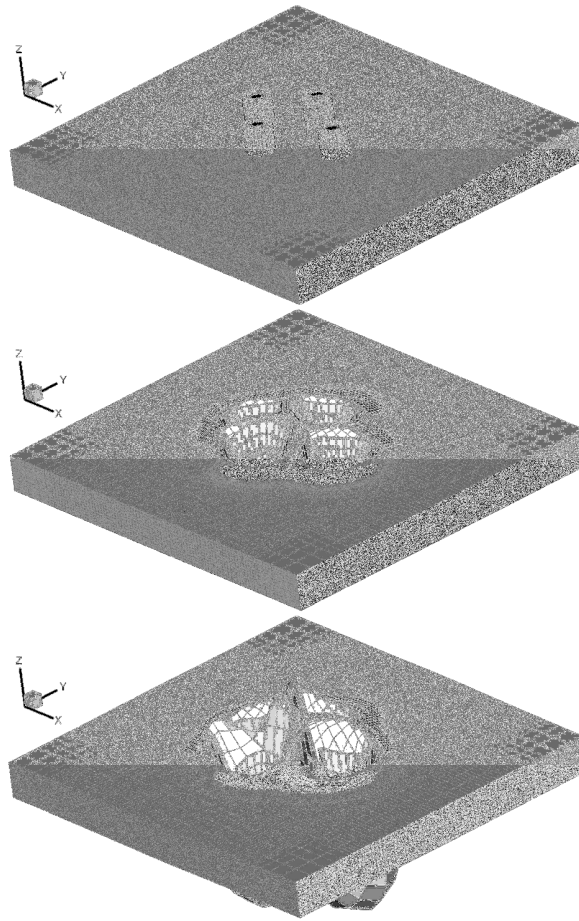


Fig. 4. Configurations of a group of four impactors at various instants of time: 0, 8, 20 μ s

Expansion-caused destruction was simulated using a kinetic model of active-type spall damage, which describes the growth and collapse of pores, processes that determine the time evolution of the properties of the material and the relaxation of stresses. The criterion of shear-induced destruction, which occurs in the region of intense interaction and deformation of impactors brought in contact, was the critical specific energy of shear deformation. The pressure in an undamaged substance is a function of the specific volume and specific internal energy; it is calculated using a Mie–Grüneisen-type equation of state, the coefficients of which are determined from the parameters of the Hugoniot curve.

The governing equations establish a relationship between the components of the stress deviator tensor and the deformation rate tensor and include the Yauman derivative. To describe the characteristics of plastic deformations, we used the Mises yield criterion. The strength characteristics of the medium (the shear modulus and dynamic yield point) depend on the temperature and extent of damage of the material.

We considered the problem of interaction of a group of identical cylindrical impactors (two to four) with a finite-thickness target. The centre-of-mass velocity vectors of the impactors before interaction coincided with their symmetry axes and were directed perpendicularly or at an angle to the face surface of the target. The initial conditions ensured that no internal stresses existed in the sample before interaction and that no forces were applied to the free surfaces; the conditions of sliding were assumed to exist at the interface between the impactors and the target. The finite-element relationships were used to solve the problem [7].

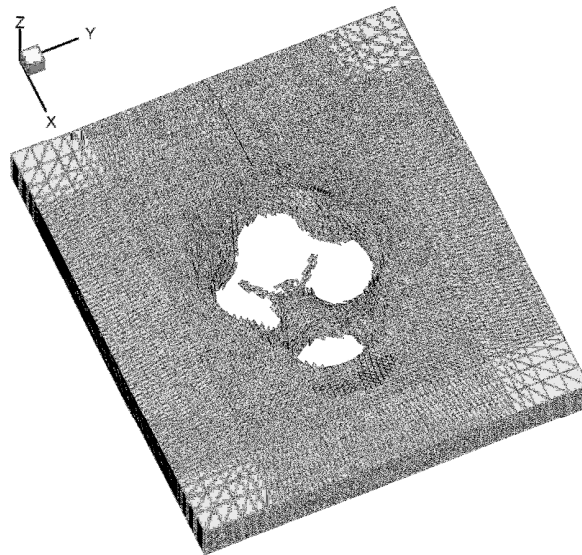


Fig. 5. Numerically simulated shape of the rear side of the target after the interaction with a group of four impactors

We simulated the interaction of compact cylindrical impactors, 6 mm in diameter and in height, with an 8-mm-thick target. The initial velocity of the impactor was 2690 m/s. The impactors and target were made of steel. Figure 4 shows typical chronograms of the penetration of a group of four impactors into the target incident along the normal to the surface.

Fig. 5 shows the computed shape of the rear side of the target for the problem under consideration. In this case, the initial distances between the impactors were such that the interaction with a group of four impactors produced a nearly elliptic rear spall region common for all the four impactors. The three through holes in the target produced by the three impactors closest to each other merge, while the hole produced by the fourth (more remote) impactor remains isolated.

Table

Initial distance between the impactors	Incidence angle	Result of interaction
$2d_0$	$0^\circ-30^\circ$	Combined region of spall damages; joint through hole in the target
$1.5-3.5d_0$	60°	
$3-3.5d_0$	$0^\circ-30^\circ$	Interaction of spall regions in the target between the impactors; the formation of additional regions of defects; isolated through hole in the target
$4.5d_0$	60°	
$>4.5d_0$	$0^\circ-30^\circ$	Independent action of each impactor

Results of interaction of a group of impactors with a target

The numerical simulation of concluding stages of the process of high-velocity interaction, accompanied by considerable plastic deformations and destruction, becomes necessary when one describes the impact of a group of particles taking place at different times, the initial stage of contact of the last particle with the barrier being coincident with concluding stages of interaction for the first particle. As criteria of complete destruction we have used in this paper the critical value of the specific volume of micro damages and the critical value of the specific energy of shear deformations. The behaviour of destructed material was not further simulated in numerical calculations. The results obtained are in close agreement with the relevant experimental data (fig. 3).

Results of interaction of a group of impactors with a target are shown in the table. The results obtained can be used to formulate simple criteria for estimating the degree of mutual influence of the impactors and for elucidating the character of target destruction.

Conclusion

The developed procedures and facilities are well suited for the laboratory simulation of the launching and the following action of a fragment debris (with a given distribution of particles in the flow over front and depth) on objects of space equipment.

The final stages of the formation of through holes were simulated with the use of complete destruction criteria. Criteria for estimating the mutual influence of the impactors and the character of destruction of the target at incidence angles of from 0° to 60° and distances between the impactors of from 1.5 to 5.0 impactor diameters were obtained.

There are difficulties in image analysis of both experimental investigations and numerical simulations. In the first case the problem is in the high rate of the process, in the latter case the problem is in 3D computations and a lot of numerical nodes.

This work was supported by the Russian Foundation for Basic Research (project no. 13-08-00509).

References

1. Experimental and theoretical investigation of encounter of a group of particles with protecting elements of spacecraft / S.A. Zelepugin [et al.] // *Cosmic Research*. – 2008. – Vol. 46, no. 6. – P. 529–539.

2. Launching and impact of a group of high-velocity bodies / I.E. Khorev [et al.] // *Doklady Physics*. – 2003. – Vol. 48, no. 3. – P. 146–150.

3. On the mechanism of collective action of a flux of solid particles upon an obstacle / E.E. Lin [et al.] // *Technical Physics Letters*. – 2002. – Vol. 28, no. 9. – P. 750–751.

4. Breaking obstacles by a group of high-velocity bodies / I.E. Khorev [et al.] // *Doklady Physics*. – 1999. – Vol. 44, no. 12. – P. 818–821.

5. Zelepugin, S.A. Modeling of the destruction of targets during a high-velocity impact / S.A. Zelepugin, A.S. Zelepugin // *Russian Journal of Physical Chemistry B*. – 2008. – Vol. 2, no. 2. – P. 246–250.

6. Zelepugin, S.A. Numerical modeling of sulfur – aluminum interaction under shock-wave loading / S.A. Zelepugin, V.B. Nikulichev // *Combustion, Explosion and Shock Waves*. – 2000. – Vol. 36, no. 6. – P. 845–850.

7. Effect of discretization in calculating three-dimensional problems of high-velocity impact by the finite-element method / V.A. Gorelski [et al.] // *Computational Mathematics and Mathematical Physics*. – 1997. – Vol. 37, no. 6. – P. 722–730.

STUDY OF DIGITAL FILTER BANKS AND THEIR SOFTWARE-HARDWARE IMPLEMENTATIONS FOR WIDEBAND MONITORING

D.M. Klionskiy, D.I. Kaplun, A.S. Voznesenskiy, V.V. Gulvanskiy
Saint-Petersburg Electrotechnical University “LETI”, Russia

e-mail: klio2003@list.ru, mitya_kapl@front.ru, a-voznensenskiy@yandex.ru, slava-a-a@mail.ru

The present paper discusses digital filter banks in the context of wideband radio monitoring tasks including DFT-modulated filter banks and the weighted overlap-add (WOLA) algorithm. Filter bank software-hardware implementations are considered on the basis of Central Processing Unit (CPU) and Compute Unified Device Architecture (CUDA) with the use of Graphics Processing Unit (GPU). The main filter bank forms are studied and compared by their execution time in MATLAB. It is shown that CUDA technology is efficient for processing large datasets and outperforms computation results on CPU.

Introduction

The term “monitoring” is determined as a set of operations for systematic and continual information collection about various parameters of a complex object or process [1–3]. Monitoring usually involves tracking condition of the main system parameters and deviation search in these parameters for subsequent information processing and decision-making.

There are many crucial applications of wideband monitoring. Related tasks are encountered in vibration processing, radio wave propagation, hydroacoustics, biomedicine, geophysics, ecology, etc. In the paper we will focus our attention on *radio monitoring* and signal processing by digital filter banks (henceforth referred to as filter banks).

1. Radio monitoring

At present *radio monitoring* is frequently applied to communication control, radar, information interception and protection, target detection, etc., which are conducted by civil and military departments in real time in a wide frequency band (as a rule, extending from zero to several GHz [3]). On the air there are usually a large number of various radiations to be traced and registered in real-time with the least possible energy losses.

Radio monitoring includes signal detection, signal denoising (noise removal), signal classification, direction-finding, estimation of the carrier frequency, phase, delay, etc.

The most substantial requirements to existing radio monitoring systems are listed below:

- 1) accurate mathematical and software tools for signal detection and direction-finding;
- 2) anti-interference hardware-software tools with low hardware costs for signal detection, classification, demodulation, and decoding.

The main requirement, though, is the ability to function properly in real-time. This inevitably leads to increasing the rate and extent of calculations and, hence, dramatic rise in hardware costs and computational resources, which might prove to be very expensive. Such hardware and software complexes are based on digital signal processing and, in particular, *digital filtering* techniques.

Monitoring is usually performed by multichannel systems [4–6]. One possible solution is the application of the discrete Fourier transform (DFT) [6, 7]. However, modulations and their combinations (e.g., amplitude and frequency modulations) often make DFT-based techniques unsatisfactory for further signal analysis [6]. For this reason *filter banks* were offered [4–6] – parallel structures for multichannel signal processing. Filter banks have many parallel channels and allow us to perform time-frequency analysis of all frequency bands simultaneously.

2. Filter banks. DFT-modulated filter banks

A DFT-modulated filter bank (including analysis filter bank and synthesis filter bank) [4–6] is based on one *low-frequency prototype filter (LF-prototype filter)* with finite impulse response (FIR-filter) and passband defining the bandwidth of each channel. This filter can be designed with the help of conventional techniques such as the windowing and Chebyshev methods [7–10]. *Sub-band processing* (in each frequency band) normally includes denoising, demodulation, amplification, filtering, spectral analysis, time-frequency analysis, etc.

Consider MATLAB simulation of a DFT-modulated filter bank with the following parameters: *number of channels*: $K = 32$; *sampling rate*: $F_s = 100$ kHz; *channel bandwidth* (two-sided): $\Delta f = 3125$ Hz (this bandwidth is nearly the same as the telephone channel bandwidth). There are three main implementations of uniform filter banks [6] explored in the paper: *direct form with full modulation (direct form)*; *critically decimated polyphase form*, and *resampled polyphase form*.

LF-prototype filter was designed by the windowing method with Kaiser window and the following parameters: *filter length*: $N_h = 6245$; *Kaiser window parameter*: $\beta = 12.25$; *bandwidth* (one-sided): 1563 Hz; *shape factor*: 1.24; *signal suppression (on passband edge)*: 1 dB; *signal suppression (on stopband edge)*: 120 dB; *passband flatness*: 0.5 dB.

Experimental research was based on processing a multiharmonic signal $x(n)$ consisting of 8 (eight) harmonics, whose amplitudes A_i and frequencies f_i (in Hz) change as $A_i = (i + 1)$; $f_i = 250(i + 1)$, $i = 0, 1, \dots, 7$. Signal length is $N = 10^5$ and signal-to-noise ratio (SNR) is 5 dB. The model of $x(n)$ can be presented as

$$x(n) = \sum_{i=0}^7 A_i \cos[2\pi \cdot 250(i+1)n] + e(n), \quad n = \overline{0, N-1}, \quad (1)$$

where $e(n)$ is additive Gaussian noise. Model (1) corresponds to real signals occurring in radio monitoring tasks and can therefore be used for experiments. In fig.1 there are plots of DFT computed for sub-band signals in channel 1 (fig.1, on the left – the interval $[0 \dots 1600]$ Hz is shown) and channel 2 (fig. 1, on the right – the interval $[1500 \dots 5000]$ Hz is shown), respectively.

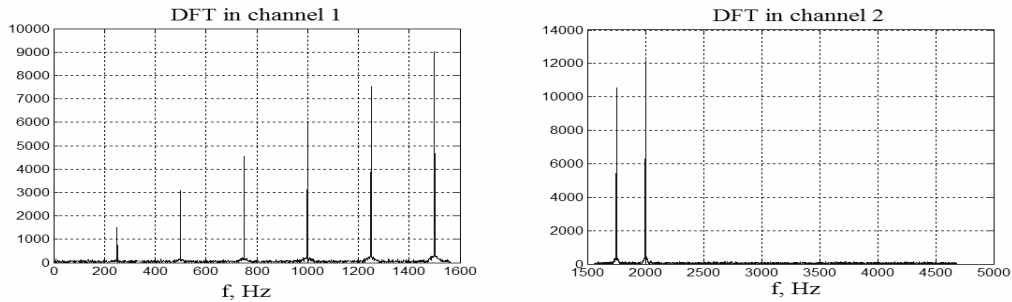


Fig. 1. Illustration of signal decomposition by a DFT-modulated filter bank (SNR = 5 dB)

As can be seen from fig. 1, the designed DFT-modulated filter bank makes it possible to separate harmonics in two adjacent channels for subsequent detection by thresholding or other techniques.

3. WOLA algorithm

A high-order LF-prototype filter is not effective since linear convolution is performed in each channel when the sampling rate is large (due to *wideband* monitoring), which results in high computational costs for real-time signal processing. Therefore it is necessary to reduce the sampling rate, for instance, by choosing a polyphase structure [6].

One of the ways of implementing a polyphase filter bank is the WOLA algorithm [6]. WOLA [6] implements a polyphase filter bank in terms of *block-by-block analysis* (the original signal is split into blocks) and WOLA is identical to the polyphase filter bank in the case of critical decimation $M = K$, where M is the decimating factor of the input signal $x(n)$.

WOLA-analysis consists of several steps:

1) weighting the m -th block by the analysis window: $y_m(r) = h(-r)x(r + mM)$, where $h(n)$ is the impulse response of a LF-prototype filter;

2) splitting the block $y_m(r)$ into overlapping segments of length K and accumulating these segments in order to find $x_m(r)$: $x_m(r) = \sum_{l=-\infty}^{\infty} y_m(r + lK)$;

3) calculation of $\hat{X}_k(m) = DFT\{x_m(r)\}$ and $X_k(m) = \hat{X}_k(m)W_k^{-kmM}$.

For the next block ($m+1$) the window is shifted by M samples and a new segment of length N_h is processed. The synthesis algorithm (*WOLA-synthesis*) is inverse to *WOLA-analysis* [6].

In fig. 2 there is a plot showing the false-alarm probability and the correct detection probability versus SNR.

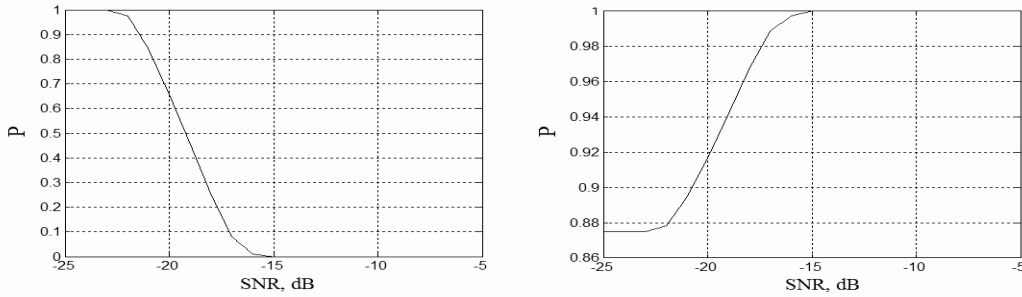


Fig. 2. False-alarm probability (on the left) and correct detection probability (on the right) versus SNR

The analysis of fig. 2 illustrates that for $SNR \geq -15$ dB the false-alarm probability tends to 0 (zero) and the correct detection probability tends to 1 (one). Smaller SNR values ($SNR < -15$ dB) lead to unsatisfactory signal detection results in filter bank channels.

4. Software-hardware implementation of a DFT-modulated filter bank and WOLA

The main factor of filter bank implementation that affects computational complexity is a LF-prototype filter. LF-prototype order mainly depends on its shape factor, passband flatness, and channel bandwidth. Thus, it is necessary to design effective FIR-filters using the minimum computational complexity criterion for reducing hardware and software costs and accelerating signal analysis by filter banks (various implementation forms).

Software implementations of a DFT-modulated filter bank and the WOLA-algorithm have been performed in MATLAB using Signal Processing Toolbox, Statistics Toolbox, and MATLAB Profiler (for estimating the execution time for all the algorithms under study).

Experiments were done on a computer with *CPU* - Intel Core i7-3630QM 2.4 GHz; *RAM* – DDR3 16 GB; *OS* – Windows 7 64 bits; *video card* - NVidia GeForce GT650M (384 core GPU, graphic base frequency 850 MHz, and video card memory 2 Gb). Below there are some examples showing the execution time for different implementation forms ($N = 10^5$, $M = K = 32$, and $N_h = 6245$). The same signal (1) was used as in sections 2 and 3.

1) The *direct implementation form* takes nearly 15 sec. for an analysis filter bank and 21 sec. for a synthesis filter bank, which might not satisfy real-time requirements.

In order to reduce the execution time it is possible to employ MATLAB functions that execute fast convolution (for time-consuming digital filtering) using the fast Fourier transform. Application of such functions has led to the following results: nearly 0.68 sec. – for an analysis filter bank and 0.75 sec. – for a synthesis filter bank, which is much better for real-time computations.

Time costs can be reduced more significantly if we use parallel computing by occupying several CPU cores (1 core is used by default, but we distributed computations among 4 cores). When the number of channels is large ($K > 300$) and long datasets are processed ($N > 10^6$), parallel computing makes computations nearly 3.8 times faster. However, if the number of channels is smaller, parallel computing does not always lead to such a substantial gain.

2) The *polyphase implementation form* has proved to be the fastest with the execution time equal to nearly 0.008 sec. for an analysis filter bank and 0.008 sec. for a synthesis filter bank (for critical decimation the lengths of polyphase filters and filtered signals are K times smaller than N).

3) As regards *WOLA*, nearly 0.18 sec. were needed for WOLA-analysis and 1.25 sec. – for WOLA-synthesis. WOLA-synthesis operates slightly longer than WOLA-analysis because it requires periodic extension [6] during its fulfilment. WOLA running time mostly depends on N_h and this time normally exceeds that for a polyphase filter bank.

In fig. 3 there are results of execution time estimation for different N (from $N = 10^5$ to $N = 10^6$) and different implementation forms (direct form, polyphase form, and WOLA algorithm).

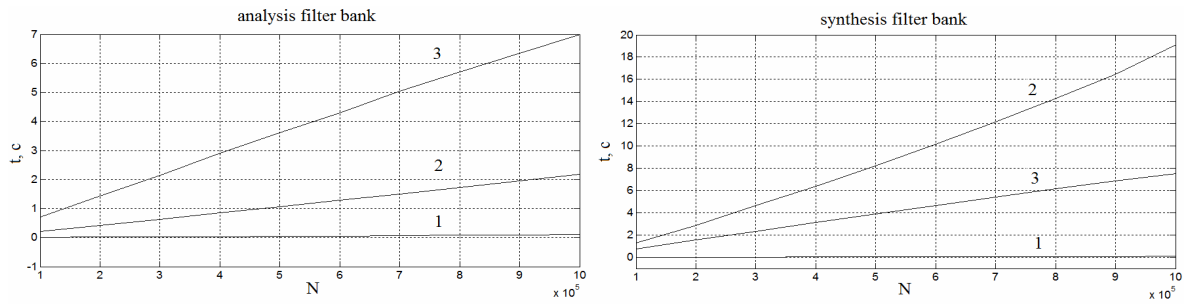


Fig. 3. Execution time for an analysis filter bank (on the left) and a synthesis filter bank (on the right) versus N (1 – polyphase implementation form; 2 – WOLA algorithm; 3 – direct implementation form)

Thus, the results for WOLA are better than those for the direct implementation form, but still yield to those provided by the polyphase implementation form (for an *analysis filter bank*, which is applied more often in radio monitoring tasks than a synthesis filter bank).

Filter banks are directed at parallel signal processing and therefore it is most reasonable to choose parallel structures for hardware implementation. Among such structures are field-programmable gate arrays (FPGA) and computation devices based on CUDA [11]. CUDA implies using GPU for handling non-graphical tasks.

The outcome of the filter bank implementations is given in tables 1-3. The *first* column indicates data size (in samples and Mbytes), the *second* column contains execution time for GPU without data transfer (from RAM to video card memory and the other way round); the *third* column contains the same information as the second column, but includes time for data transfer (from RAM to video card memory and the other way round); the *fourth* column indicates execution time for CPU.

Table 1

Execution time using CPU and CUDA for the direct implementation form

Data size (samples / Mb)	CUDA without data transfer (time), ms	CUDA with data transfer (time), ms	CPU (time), ms
3000000 / 11	1443	1658	260542
8000000 / 30	3759	4334	680000
16000000 / 61	6752	7902	1360000
30000000 / 114	11937	14093	2550000

Table 2

Execution time using CPU and CUDA for the critically decimated polyphase implementation form and WOLA-algorithm

Data size (samples / Mb)	CUDA without data transfer (time), ms	CUDA with data transfer (time), ms	CPU (time), ms
3000000 / 11	293	508	46931
8000000 / 30	320	895	122656
16000000 / 61	346	1496	246834
30000000 / 114	437	2593	459354

Table 3

Execution time using CPU and CUDA for the resampled polyphase implementation form

Data size (samples / Mb)	CUDA without data transfer (time), ms	CUDA with data transfer (time), ms	CPU (time), ms
3000000 / 11	388	603	80406
8000000 / 30	492	1067	215469
16000000 / 61	721	1871	435744
30000000 / 114	1102	3258	817088

Detailed analysis of tables 1-3 leads to the following conclusions:

- application of CUDA provides significant reduction in execution time (reduction in computational costs) in comparison with CPU;
- polyphase implementation form is the least time-consuming, whereas direct implementation form is the most time-consuming among the three different forms studied in the paper;
- increasing data size leads to a rise in the difference between execution times for CUDA and CPU;

–data transfer time is nearly half of total processing time for CUDA, which means that further improvements of CUDA have to be developed and tested;

–CUDA makes it possible to reduce hardware costs and implement more complicated computational algorithms of signal processing.

Conclusion

We have considered different forms of a DFT-modulated filter bank and the WOLA algorithm in wideband radio monitoring tasks. Filter bank implementations were compared and CUDA made it possible to dramatically reduce processing time. The paper was prepared in SPbETU and is supported by the grant of the Russian Foundation for Basic Research ("My first grant") № 14-07-31250/14 and Contract № 02.G25.31.0058 dated 12.02.2013 (Board of Education of Russia).

References

1. Porcino, D. Ultra-Wideband Radio Technology: Potential and Challenges Ahead / D. Porcino, W. Hirt // IEEE Communications Magazine. – 2003. – Vol. 41 (7). – P. 66–74.
2. Hirt, W. Ultra-Wideband Radio Technology: Overview and Future Research, Comp. / W. Hirt // Computer Communications. – 2003. – Vol. 26 (1). – P. 46–52.
3. Order reduction of a low-frequency prototype filter by its magnitude response symmetrization in the task of filter bank design for wideband monitoring / D.I. Kaplun [et al.] // Proc. of the Intern. Conf. Dedicated to the 50th Anniversary of MREI-BSUIR. – Belarus, 2014. – P. 3–8.
4. Piotrowski, A. Digital filter banks: analysis, synthesis and implementation for multimedia systems / A. Piotrowski, M. Parfieniuk. – Poland : Wydawnictwo politechniki bialostockiej, 2006. – 389 p.
5. Vashkevich, M.I. Practical design of multi-channel oversampled warped cosine-modulated filter banks / M.I. Vashkevich, W. Wan, A.A. Petrovsky // Proc. Of CCWMC 2011. – China, 2011. – P. 1–6.
6. Crochiere, R.E. Multirate digital signal processing / R.E. Crochiere, L.R. Rabiner. – USA : Prentice Hall, 1983. – 411 p.
7. Mitra, S.K. Digital Signal Processing: A Computer-Based Approach / S.K. Mitra. – USA : McGraw-Hill, 1998. – 940 p.
8. Antoniou, A. Digital Filters: Analysis, Design, and Applications / A. Antoniou. – USA : McGraw-Hill, 1993. – 689 p.
9. Oppenheim, A. Digital Signal Processing / A. Oppenheim, R. Shafer. – USA : Prentice Hall, 1975. – 585 p.
10. Rabiner, L.R. Theory and applications of digital signal processing / L.R. Rabiner, B. Gold. – USA : Prentice Hall, 1975. – 762 p.
11. Kirk, D.B. Programming Massively Parallel Processors: A Hands-on Approach / D.B. Kirk, Wen-mei W. Hwu. – USA : Morgan Kaufmann, 2010. – 280 p.

ARCHITECTURE OF THE DCT-IDCT PROCESSOR FOR LOSSLESS SCHEME CODING

V. Kliuchenia, A. Petrovsky
Belarusian State University of Informatics and Radioelectronics, Minsk
e-mail: lucky_victor@rambler.ru, palex@bsuir.by

This paper is proposed a structural solution forward discrete cosine transform (DCT) and inverse (IDCT) as a single computing system module for lossless data compression scheme. Ladder structures are used for data compression lossless mode or perfect reconstruction. One step of ladder structure is consisted of blocks DCT and IDCT.

Introduction

Nowadays, portable multimedia real-time systems are widely used. This systems are used standard for video, audio and images compression, such as H.261/3/4/5, MPEG-1/2/4 and JPEG. The core of these standards is the discrete cosine transform DCT-I, -II, -III, -VIII-type. Expediency lossless image compression associated with the need unbiased information processing applications such as image processing with satellites, unmanned aerial vehicles boards , medical images. Such an encoding with lossless mode has already in JPEG2000 and HD Photo (JPEG-XR), etc. However, existing implemented technology [1, 2] such as JPEG2000, JPEG-XR can't replace the JPEG standard, because various electronics devices and software are widely supported format JPEG. This means that lossless image encoding must be compatible to the parameters and characteristics with JPEG standard.

1. Mathematical description lossless scheme coding

For perfect reconstruction will use a ladder structure are proposed in [3], one of the properties which is a full data recovery (fig. 1).

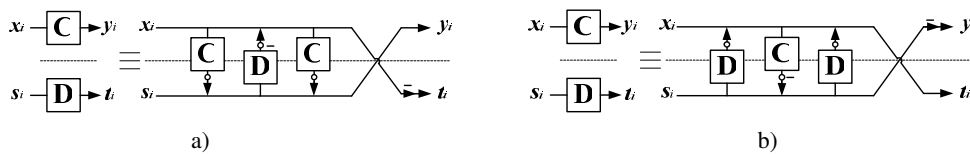


Fig. 1. Ladder structure of lossless coding (white circles represent rounding operations):
a) forward transform, b) inverse transform [3]

Mathematical description of the forward (fig. 1, a) and inverse (fig. 1, b) transform block ladder is following [3]:

$$\begin{bmatrix} C & 0 \\ 0 & D \end{bmatrix} = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} \begin{bmatrix} I & 0 \\ C & I \end{bmatrix} \begin{bmatrix} I & -D \\ 0 & I \end{bmatrix} \begin{bmatrix} I & 0 \\ C & I \end{bmatrix}; \quad (1)$$

$$\begin{bmatrix} C & 0 \\ 0 & D \end{bmatrix}^{-1} = \begin{bmatrix} D & 0 \\ 0 & C \end{bmatrix} = \begin{bmatrix} I & 0 \\ -C & I \end{bmatrix} \begin{bmatrix} I & D \\ 0 & I \end{bmatrix} \begin{bmatrix} I & 0 \\ -C & I \end{bmatrix} \begin{bmatrix} 0 & -I \\ I & 0 \end{bmatrix}. \quad (2)$$

Forward and inverse DCT, which are described as DCT-II and DCT-III type respectively, calculated using the following formulas:

$$[C]_{m,n} = \sqrt{\frac{2}{M}} \cdot c_m \cos\left(\frac{m(n+\frac{1}{2})\pi}{M}\right), [D]_{m,n} = \sqrt{\frac{2}{M}} \cdot c_n \cos\left(\frac{n(m+\frac{1}{2})\pi}{M}\right), \quad (3)$$

where $D = C^{-1} = C^T, 0 \leq m, n \leq M - 1$, $[C]_{m,n}$ – matrix forward DCT or DCT-II type, $[D]_{m,n}$ – matrix inverse DCT or DCT-III type, m – number of columns, n – number of rows, $M = 2^n (n \in N)$.

2. Hardware module of calculating DCT-IDCT

This system (fig. 1) is acceptable to use any existing implementation of the DCT image encoding. In order to implement this system (fig. 1) need to connect three consecutive DCT processor that requires a lot of hardware costs FPGA. Therefore, it is advisable to combine the DCT and IDCT in one module. The implementation of DCT blocks and IDCT for lossless scheme is proposed to use the universal module (fig. 2), which allows us to compute the DCT and IDCT either simultaneously or sequentially in different modes.

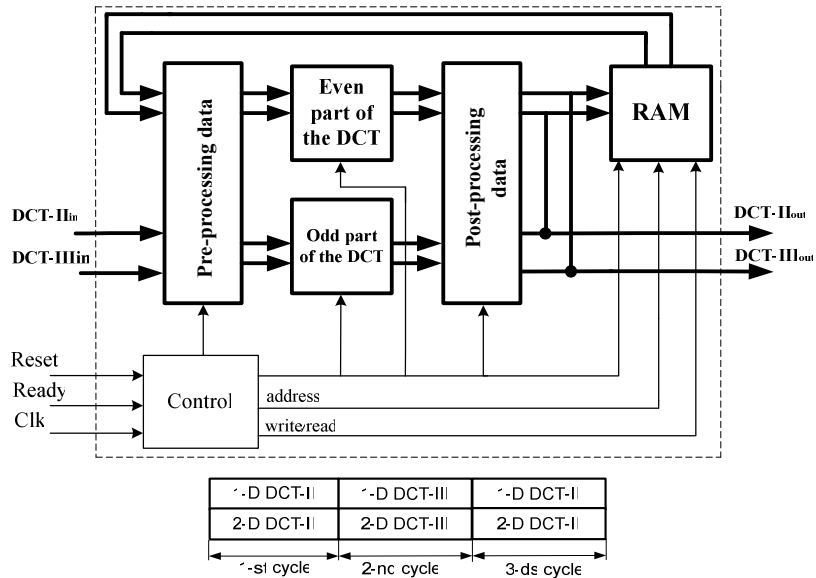


Fig. 2. Architecture of module calculating DCT-IDCT

Encode/decode images are used 2-D (two-dimensional) DCT dimension 8x8. Calculation is made first with respect to the rows (1-D (one-dimensional) DCT), while maintaining the intermediate results in the memory (RAM), and then by column (2-D DCT) as for the forward and inverse DCT. In lossless mode calculation is done according to the scheme 2-D DCT /2-D IDCT/2-D DCT (Fig. 2) in three cycles. Vector of input and output values is as follows, respectively:

$$\mathbf{x} = [x_0 \ x_1 \ x_2 \ x_3 \ x_4 \ x_5 \ x_6 \ x_7]^T, \mathbf{z} = [z_0 \ z_1 \ z_2 \ z_3 \ z_4 \ z_5 \ z_6 \ z_7]^T. \quad (4)$$

Using the symmetry property of DCT, prepare the data according to the formula:

$$\mathbf{a} = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} x_0 + x_7 \\ x_1 + x_6 \\ x_2 + x_5 \\ x_3 + x_4 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} x_0 - x_7 \\ x_1 - x_6 \\ x_2 - x_5 \\ x_3 - x_4 \end{bmatrix}, \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{bmatrix} = \frac{1}{2}(\mathbf{a} + \mathbf{b}), \begin{bmatrix} x_7 \\ x_6 \\ x_5 \\ x_4 \end{bmatrix} = \frac{1}{2}(\mathbf{a} - \mathbf{b}). \quad (5)$$

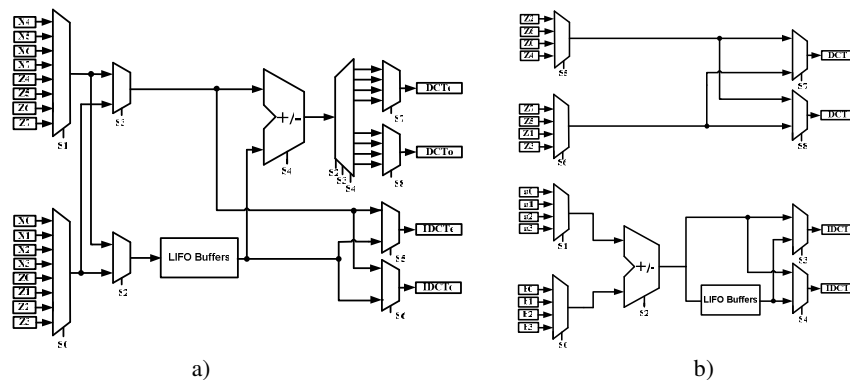


Fig. 3. Architecture of a) data pre-processing, b) data post-processing

Data pre-processing module are used (fig. 3, a) for calculating values of the vectors \mathbf{a} and \mathbf{b} . Input values are provided through multiplexers to the buffer memory, further through the adder-subtractor data

receives to the demultiplexer, where they are divided into even and odd parts forward DCT. Also input to this module already is served the calculated values from the previous cycle Z0-Z7, that accumulate in the buffer (LIFO Buffers) and also further broken down into even and odd parts (fig. 3, a). A similar situation occurs in the post-processing module. Values of the vectors \mathbf{a} and \mathbf{b} are summed or subtracted according to the formula (5), accumulated in the buffer, re-formed in order and the resulting values of \mathbf{x} comes into RAM for further calculations, and Z0-Z7 reordered in the desired sequence of output values (fig. 3, b).

3. Module of even and odd parts of DCT-IDCT based on distributed arithmetic

After pre-processing, the data arrives in calculation module of the even and odd DCT and IDCT.

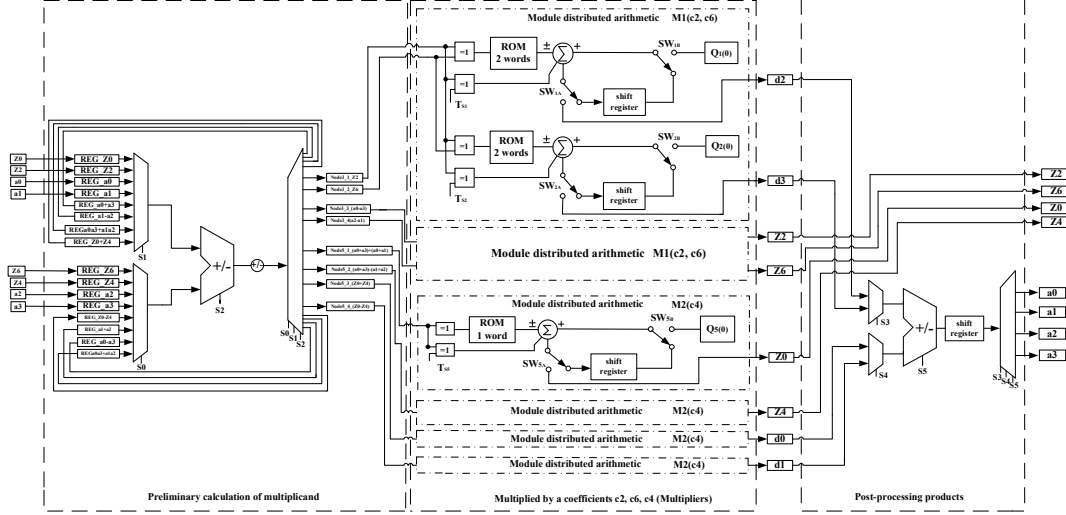


Fig. 4. Architecture of the module even DCT-IDCT

Even part of the direct and inverse DCT is expressed by the formulas:

$$\begin{bmatrix} Z_0 \\ Z_4 \\ d_0 \\ d_1 \end{bmatrix} = \begin{bmatrix} a_0 + a_2 & a_1 + a_2 \\ a_0 + a_2 & -(a_1 + a_2) \\ Z_0 & Z_4 \\ Z_0 & -Z_4 \end{bmatrix} \begin{bmatrix} c_4 \\ c_4 \end{bmatrix} = c_4 \begin{bmatrix} a_0 + a_2 + a_1 + a_2 \\ a_0 + a_2 - (a_1 + a_2) \\ Z_0 + Z_4 \\ Z_0 - Z_4 \end{bmatrix}, \quad (6)$$

$$\begin{bmatrix} d_2 \\ d_3 \\ Z_2 \\ Z_6 \end{bmatrix} = \begin{bmatrix} Z_2 & Z_6 \\ -Z_6 & Z_2 \\ a_0 - a_2 & a_1 - a_2 \\ -(a_1 - a_2) & a_0 - a_2 \end{bmatrix} \begin{bmatrix} c_2 \\ c_6 \end{bmatrix}, \quad \begin{bmatrix} d_0 \\ d_2 \\ d_3 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} a_0 + a_2 \\ a_1 + a_2 \\ a_0 - a_2 \\ a_1 - a_2 \end{bmatrix}, \quad \frac{1}{2} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} d_0 + d_2 \\ d_1 + d_3 \\ d_1 - d_3 \\ d_0 - d_2 \end{bmatrix}, \quad (7)$$

where $c_i = \cos(i\pi/16)$ – constant coefficients of DCT and IDCT. The structure of the even part of the DCT-IDCT is presented fig. 4, which consists of three parts according to the formulas (6, 7):

- pre-computation values of multiplicand, which consists of an adder-subtractor and two sets of registers where pre-calculated values or input data are saved;

- multiplication by coefficients c2, c4 and c6 are implemented using distributed arithmetic method. Since the need to obtain perfect reconstruction scheme is used where the data is processed one bit at a time (1-BAAT) [4], preliminary values of the sums coefficients c2 and c6, and coefficient c4 are stored in memory (ROM);

- the third part of the post-processing products is implemented according to formula (7).

The odd part of the forward and inverse DCT is expressed by the formulas:

$$\begin{bmatrix} Z_7 \\ Z_5 \\ Z_1 \\ Z_3 \end{bmatrix} = \begin{bmatrix} -b_3 & b_2 & b_0 & -b_1 \\ -b_1 & b_3 & b_2 & b_0 \\ b_0 & b_1 & b_3 & b_2 \\ -b_2 & b_0 & -b_1 & -b_3 \end{bmatrix} \begin{bmatrix} c_1 \\ c_3 \\ c_7 \\ c_5 \end{bmatrix}, \quad \frac{1}{2} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} Z_1 & Z_3 & Z_7 & Z_5 \\ -Z_5 & Z_1 & -Z_3 & -Z_7 \\ -Z_7 & Z_5 & Z_1 & -Z_3 \\ -Z_3 & Z_7 & Z_5 & Z_1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_3 \\ c_7 \\ c_5 \end{bmatrix}. \quad (8)$$

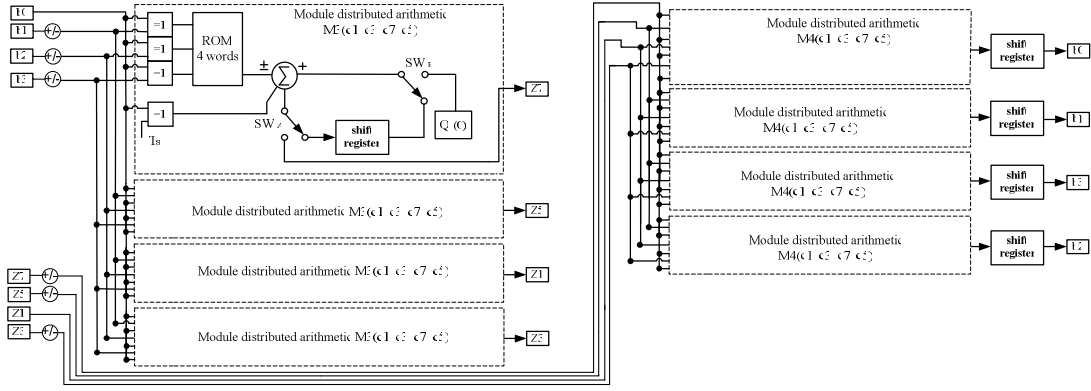


Fig. 5. Architecture of the odd module DCT-IDCT

The structure of the odd part of the DCT is presented in figure 5 which consists of a distributed arithmetic modules, all possible values of multipliers are presented as a sum of coefficients c_1 , c_3 , c_7 and c_5 , which is stored in the ROM. Calculations are made according to the formula (8).

4. Experimental results

This implementation are compared with existing ones. Since DCT architecture processors, which are taken for comparison, were synthesized on FPGA Spartan-2, then for a more accurate comparison, we using VHDL-descriptions synthesized calculation module DCT-IDCT under this architecture FPGA for two-dimensional DCT-IDCT module (table).

Comparison of existing implementations two-dimensional DCT and DCT-IDCT on Xilinx FPGA Spartan-2 XC2VP30

Architecture	[5]	[6]	[7]	[8]	[9]	This work
Functions	DCT	DCT	DCT	DCT	DCT/IDCT	DCT/IDCT
# of 4 input LUTs	2990	990	10310	2618	2237	1109
# of Slices	1872	612	5729	2823	1352	627
# of Slice Flip Flops	1837	652	3736	3431	1170	887
Clock Frequency	99	110	149	107	168	120
Max. Delay (ns)	10.1	9.1	6.7	9.3	6.2	8.3

If you look at table, we see that the proposed implementation requires twice less hardware costs compared to similar to that described in [9]. At some times takes less FPGA chip area compared to [5], [7] and [8], which shows only one implementation of the DCT.

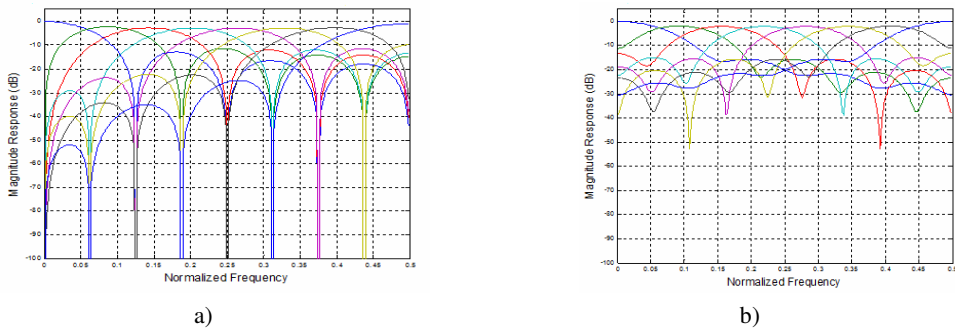


Fig. 6. Frequency response: a) DCT-II; b) DCT-III

The disadvantage of ladder structure of lossless coding (fig. 1) is absent of regularity first order property in frequency response IDCT (fig. 6, b). This causes to the effect of DC leakages (DCT all filters have a constant component), so when decoding image reconstruction is not perfect. This effect is manifested as "chessboard" (fig. 7, b), when a hardware implementation on a fixed-point arithmetic. For

encode a lossless data entered additional information block S_i dimension $N \times N$ as block DCT. Initially S_0 is the zero matrix, and S_i is iterative transformation from S_0 as

$$s_i = Ds_{i-1} \text{ for } i = 1, 2, \dots, n-1. \quad (9)$$

As a result, the implementation of lossless image coding/decoding will be represented by the formulas

$$\begin{bmatrix} y_0 \\ \vdots \\ y_{n-1} \\ s_n \end{bmatrix} = \begin{bmatrix} C & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & C & D^n \end{bmatrix} \begin{bmatrix} x_0 \\ \vdots \\ x_{n-1} \\ s_0 \end{bmatrix}, \begin{bmatrix} x_0 \\ \vdots \\ x_{n-1} \\ s_0 \end{bmatrix} = \begin{bmatrix} D & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & D & C^n \end{bmatrix} \begin{bmatrix} y_0 \\ \vdots \\ y_{n-1} \\ s_n \end{bmatrix}, \quad (10)$$

where s_n also encoded with all y_n , C and D is the matrix of direct and inverse DCT according to formulas (3). Should pay attention what $s_i \neq 0$ due to rounding errors on each ladder step.

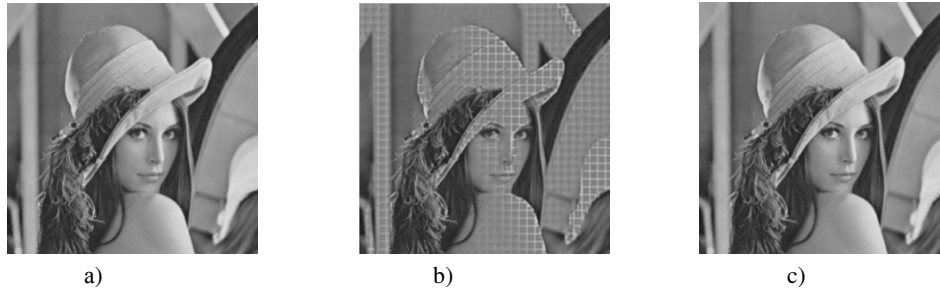


Fig. 7. Results image compression of Lena 512x512: a) original image; b) decoded image without s_n ; c) decoded image with information block s_n

Conclusion

In this article are presented the implementation of block ladder structure for encoding data with lossless mode as a single universal module, which calculates a two-dimensional forward and inverse discrete cosine transforms in data compression modes with lossless for perfect reconstruction or the full data recovery. This structure allows to save space on the FPGA chip when implementing the algorithm for encoding and decoding images, which is very important for embedded multimedia systems.

References

1. Low-complexity hierarchical lapped transform for lossy-to-lossless image coding in JPEG XR / HD Photo / C. Tu [et al.] // Proceedings of the SPIE Applications of Digital Image Processing XXXI. – San Diego, USA, 2008.
2. Suzuki, T. Realization of lossless-to-lossy image coding compatible with JPEG standard by direct-lifting of DCT-IDCT / T. Suzuki, M. Ikehara // Proceedings of 2010 IEEE 17th Intern. Conference on Image Processing. – Hong Kong, 2010.
3. Suzuki, T. Integer DCT Based on Direct-Lifting of DCT-IDCT for Lossless-to-Lossy Image Coding / T. Suzuki, M. Ikehara // IEEE Transactions on Image Processing. – 2010. – Vol. 19, no. 112. – P. 2958–2965.
4. Ключеня, В.В. Выбор оптимальной реализации структурного решения процессора ДКП на распределенной арифметике / В. В. Ключеня, А. А. Петровский // Доклады БГУИР. – 2010. – № 7. – С. 66–75.
5. Chen, Y.H. High throughput DA-based DCT with high accuracy error-compensated adder tree / Y.H. Chen, T.Y. Chang, C.Y. Li // IEEE Trans. VLSI Syst. – 2011. – Vol. 19, no. 4. – P. 709–714.
6. Chen, Y.H. A high performance video transform engine by using space-time scheduling strategy / Y.H. Chen, T.Y. Chang // IEEE Trans. VLSI Syst. – 2012. – Vol. 20, no. 4. – P. 655–664.
7. Sun, C.C. Low-complexity multi-purpose IP core for quantized discrete cosine and integer transform / C.C. Sun, P. Donner, J. Gotze // International Symposium on Circuits and Systems (ISCAS 2009). – Taipei, Taiwan, 2009.
8. A pipelined fast 2D-DCT accelerator for FPGA-based SoCs / A. Tumeo [et al.] // IEEE Computer Society Annual Symposium (VLSI 2007). – Porto Alegre, Brazil, 2007.
9. A High-Throughput and Area-Efficient Video Transform Core with a Time Division Strategy / Y.-H. Chen [et al.] // IEEE Transactions on Very Large Scale Integration (VLSI) Systems. – 2014. [In print]

AN AGENT-BASED MODEL OF GERMAN-FRENCH INTER-BORDER RESIDENTIAL MOBILITY IN STRASBOURG REGION

A. Kovalev

Hanse Ventures, Hamburg, Germany
e-mail: alexei.kovalev.germany@gmail.com

The purpose of this paper is to present a first version of an agent-based residential mobility model which is closely related to the job-associated, inter-EU migration and commuting problems. At this stage the model provides the basic functionality such as modeling of how the property market agents (potential householders) would behave in French-German border region while looking for a property to be purchased. Then the model is extended further by way of adding property selling functionality. The model execution resulted in several outputs and statistics which represent migration processes under different conditions.

Introduction

Recently, the mobility phenomena become one of the major issues for both scientific research and real practice in Europe [1]. The job-related commuting and residential mobility problems are known to be closely related to each other [2] and typically originated from the bordering regions such as Luxembourg-France, France-Switzerland, Denmark-Germany, and Germany-France [1, 3].

It is known that for a number of economics, social, and financial problems the agent-based behavioral simulation models hold various advantages over the other approaches [4]. They include the high flexibility, capability of modeling complex spatial systems, correspondence of modeled subjects to real-life agents, etc. Also, it is proven that agent-based simulation approach still works where math and statistics does not.

This paper presents an agent-based model of German-French inter-border residential mobility in Strasbourg region with such specific characteristics as the population density variation in the area being considered. The model is implemented using the NetLogo language and modeling environment [5] under Windows operating system.

1. General Structure, Functions, and User Interface of the Model

There are two versions of the simulation model. The first (basic) one represents how the property market agents (i.e., potential Householders) would behave in French-German border region while looking for a property to be purchased. The second version is a further extension of the first one what is done by way of adding property selling functionality. So far no in-depth mechanisms such as moving from Germany (DE) to France (FR) and vice versa in order to get new job with all related job market dynamics (e.g., types of employers, their capacity, professional and educational requirements, individual employment preferences, contract type, discharging and resignation rules, etc.) have been implemented so far because it is out of scope of the preliminary study presented in this paper. Same goes for the fact that despite the model execution resulted in several outputs and statistics, only few of them are displayed in form of corresponding plots of user interface.

It should be noted that in the text below the “household” term is used as a sort of conditional identifier of various software objects and related notions. This is why sometimes it may disagree with its original meaning characteristic for “normal” English language. In addition, some terms denoting model’s objects are capitalized in the text by similar reasons.

The first version of the model addresses the following research question: “How the Property Seekers would behave depending on the major dwelling parameters”. The parameters are as follows.

- Basic property prices (here: random values uniformly distributed in the given range which may be different for each country, inputted in “K Euros”).
- Actual level of the Property Purchasing Tax in the country.
- Expenses associated with re-location from FR to DE and vice versa (when relevant to the specific circumstances). This value is a monetary equivalent of all the expenses, losses, and other inconveniences (e.g. living in another Language environment) associated with relocation from the home country abroad. Clearly, they may vary a lot.
- Dwelling price variations depending on the Attractiveness of its specific location (presented by an Attractiveness Map provided in the form of input PNG image file).

Fig. 1 illustrates user interface of the first version of the model and provides an example model output for given set of control parameters.



Fig. 1. User interface of the first version of the model along with the example of modeling results (right side of the panel) for given set of control parameters depicted on the left side with the help of corresponding sliders

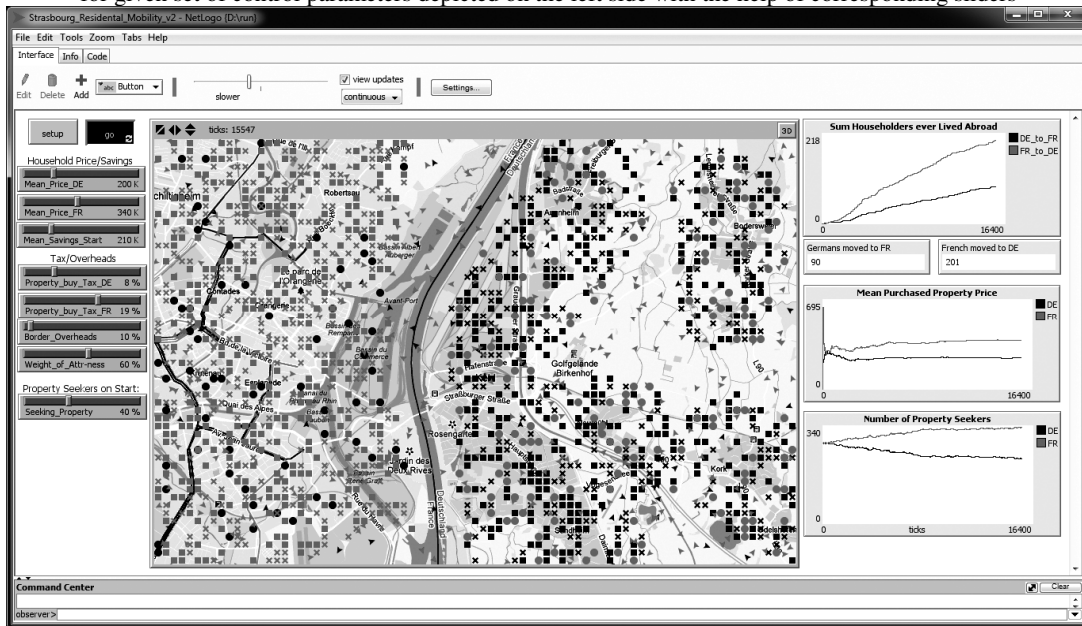


Fig. 2. User interface of the second version of model presenting similar data as the ones depicted in fig. 1

In the second version of the inter-border residential mobility model an additional Property Selling procedure is added what completes the model functionality. In contrast to the first version where the process actually stops as soon as all the Property Seekers purchasing the desired and affordable property, the modeling process here is virtually endless. This allows observing convergence of target variables to their asymptotic values or their divergence under various sets of control parameters. An illustration of the user interface and example output of the second version of the model provided in fig. 2. Note that the first output plot situated on the top-right corner of the screen represents now the total amount of householders who have ever been lived abroad for certain period of time. This is why corresponding curves are going up monotonically.

2. Major Agents and their Display Agreements

There are two main types of agents in the model:

(a) Dwellings (i.e., pieces of property or households) which characterized by the Country they are situated in, specific location, and their Attractiveness. On the model display the current Status of the Household is depicted either by a filled square meaning “occupied” or by a cross “X” meaning “remains free”.

(b) Householders who either owning a property that is “sitting” in the square-shaped property objects of the model display, or looking to purchase one, i.e. belonging to the Property Seekers who is depicted by moving “arrows”. The major characteristic of a Property Seeker is the amount of money that is dedicated Savings he/she has been accumulated for property purchase. These savings are changed according to various factors such as the previous property selling event or presence of border crossing overheads which are deducted in case the property going to be purchased in the country different from the new owner’s country of origin.

In present denotations and coloring agreement German objects are represented in black color whereas French ones are colorized by red.

The modeled geographical region is given by a section of map which can be considered here as a sort of skin placed over the NetLogo working space and which is used mostly for illustrative purposes. For current version of the model corresponding map section is borrowed from Google Maps.

Contrary to the above geographical region map, the map of Attractiveness which represents in quantitative way the attractiveness of specific property location is one of the key set of parameters of the model. This 2D map is placed onto invisible layer of working space so that each cell of working space corresponding specific property location (that is *patch* in NetLogo’s terms) can be associated with certain value of attractiveness.

3. Key Parameters, Behavioral and Communication Rules

The behavioral and communicating rules which define the model dynamics are as follows:

1. For the simplicity purposes, the number of Householders (families) and Houses (property units) is supposed to be equal. This particularly means that theoretically, each Householder may have his own separate property.

2. At the first initialization step the given fraction of potential Householders is converted to Property Seekers, status of their Households switched down to “free”, and the amount of Savings of Property Seekers is calculated as the existing one (allocated at random by given probability density distribution) plus the Price (after taxation) of the property he/she just sold.

3. Property Seekers are permanently moving over the modeled region i.e., “searching” for suitable property purchasing variant. Once a Property Seeker arrived to a free property, no matter French or German one, he/she attempts to purchase it. The purchasing decision depends on whether the amount of Savings he holds is greater than the Full Property Price or not. The Full Price is calculated using Basic Price plus all the Taxes and Overheads mentioned above.

Note that the model is designed so that the additional factors contributing to the Full Property Price including the ones lifting prices up and pushing them down, no matter whether they are deterministic or stochastic, can be easily incorporated into the model by adding few more lines into the source code of “Calculate_Full_Household_Price” procedure. However, when doing so, one should remember that consideration of greater number of factors not necessarily means better modeling.

4. If the Property Seeker has the sufficient amount of Savings the procedure of property purchasing is executed. It switches agent’s status from Property Seeker down to Householder and the amount of Savings is decreased for the full price the agent paid for the property. The property status switches from “free” to the “occupied” accordingly and the crossed square replaced by a solid one on the model’s display.

5. For better visualization of the results, in case a property item has been purchased by a Foreigner, i.e. by a French family on German territory or by a German family on French one, the property being purchased is depicted by a circle filled by Householder’s home country color rather than by the color of the country the property is situated in. This holds true until the property is sold again to a native family. Such a visualization technique eases visual monitoring of the residential mobility.

4. Implementation Details and Model Running Notes

In general, model implementation starts with a preparatory stage the main goal of which is to provide appropriate representation of all quantitative model components. These include all the formulas involved, specification of control parameters inputted by the researcher, description of stochastic input variables, which represent certain phenomena of real-life economics (typically by way of identification of certain probability density functions together with their parameters), a list of the accepted numerical values of all the bounding conditions and limits, description of the resultant modeling data, their visualization and store, etc. In some cases this stage may suppose performing some auxiliary experiments and developing particular sub-models in order to validate specific model design decisions and convince the community in the model relevance. These details are mostly omitted here due to paper size limitations. Some key points are given below, though.

The list of input parameters together with their values and ranges can be inferred from the model displays given in fig. 1 and 2. It should be noted that the general attractiveness of the region is proportional to the sum of intensities of corresponding map of attractiveness. Such factors and biases need to be considered when preparing the model runs.

A number of model parameters are not deterministic by nature and contain certain stochastic elements which are generated at random during the initialization step. Thus, different runs with the same parameter settings may produce slightly different results. In real modeling this can be easily overcome either by performing series of repeated experiments followed by statistical analysis (e.g. calculating statistics of target values) or by setting up a very large number of acting agents. The last option may require large computer resources which otherwise kept reasonably modest, namely, an ordinary PC Computer and computational time in order of tens of minutes for model convergence. The model is available for free from author and it is supplied as a package containing the following files:

- Strasbourg_Residential_Mobility_v2.nlogo (the source code of the Model itself in NetLogo language, 502 lines of code);
- Strasbourg_Map.png (a map of Strasbourg region used as a display skin);
- Strasbourg_Population_Map_81x54.png (example of Attractiveness map which at this stage has been derived from population density distribution in modeled region).

References

1. Mobility in Europe. Analysis of the 2005 Eurobarometer survey on geographical and labour market mobility / T. Vandenbrande [et al.]. – European Foundation for the Improvement of Living and Working Conditions, 2006. – 82 p.
2. van Ommeren, J. Job mobility, residential mobility and commuting: A theoretical analysis using search theory / J. van Ommeren, P. Rietveld, P. Nijkamp // *Annals of Regional Science*. – 2000. – Vol. 34 (2). – P. 213–232.
3. Buch, T. Cross-Border Commuting in the Danish-German Border Region – Integration, Institutions and Cross-Border Interaction / T. Buch, T.D. Schmidt, A. Niebuhr // *Journal of Borderlands Studies*. – 2009. – Vol. 24 (2). – P. 38–54.
4. Dieleman, F.M. Modelling residential mobility: a review of recent trends in research / F.M. Dieleman // *Journal of Housing and the Built Environment*. – 2001. – Vol. 16. – P. 249–265.
5. Railsback, S. Agent-based and individual-based modeling : A practical introduction / S. Railsback, V. Grimm. – Princeton University Press, 2011. – 352 p.

HIGH-LEVEL RENDERING PIPELINE CONSTRUCTION TOOLS

V. Krasnoproshin¹, D. Mazouka²
Belarusian State University, Minsk
e-mail: krasnoproshin@bsu.by, mazovka@bk.ru

This paper describes a methodology for the graphics pipeline extension. The introduced approach is based on specialized formal language called visualization algebra. This technique can lower visualization software development costs and build a way for further computer graphics automation.

Introduction

Computer graphics (CG) remains one of the most rich and constantly evolving fields of study in computer science. CG consists of two large parts, each studying its own problem: image recognition and image generation. Both of these problems are very broad and complex in nature. In this paper, we concentrate on image generation, i.e. visualization.

Literally every area in human life which involves computer technology requires some sort of visual representation of information. Accurate and adequate visualization becomes vitally necessary with the growing complexity of problems being solved with computers. CG provides tools and theories that target the growing requirements for visualization, but as requirements become more complex and demanding so does the need for improvement in this field.

In this paper we analyze the most widely used visualization methodology and provide a technical solution for its improvement.

1. Basic Definitions

Visualization, broadly defined, is a process of data representation in visual form. In computer graphics visualization has a special definition – **rendering**. The rendering process is implemented in computers with a set of dedicated hardware components and specialized software.

There are several rendering algorithms that lay a base for computer visualization. The most widespread are ray tracing and **rasterization** [1]. These algorithms have their own application areas: ray tracing is used for photorealistic rendering, sacrificing computation performance to physically correct output; rasterization, in its turn, is designed for high-performance dynamic applications where photorealism is not as essential.

Almost all contemporary rendering hardware implements the rasterization algorithm. A technical implementation of the algorithm is called a **graphics pipeline**. A graphics pipeline is structured as a sequence of data processing stages, most of which are programmable. These stages are responsible for data transformations specified by the rasterization algorithm's logic. The pipeline is accessible for software developers with special hardware abstraction layer libraries, available on most operating systems. The most popular libraries are DirectX (for Windows) and OpenGL (for any OS). Both of them provide a software abstraction of the underlying hardware implementation of graphics pipeline with all necessary access methods.

The graphics pipeline is a very efficient and flexible technology, but it may be difficult to use in complex applications. This is because the pipeline's instructions operate on a low level with objects like geometrical points and triangles. The best analogy here would be to compare this with the efforts of programming in an assembly language.

The problem of the pipeline's complexity gave rise to a new class of visualization software: **graphics engines**. A graphics engine is a high level interface that wraps around the pipeline's functionality, and introduces a set of tools which are much more convenient to use in real world applications. The engine is normally built around an extendible but nevertheless static computer model, which generalizes a whole spectrum of potential visualization tasks. The most popular visualization model for graphics engines nowadays is the **scene graph**.

Scene graph is not a well-defined standard model and many software developers implement it differently for different tasks. However, the implementations often have common traits, which can be summarized in the following definition. Scene graph – is a data structure for visualization algorithms, based on a hierarchical tree where every node is an object and every subnode – a part of that object. That is, a scene graph may have a node representing a chair object and direct subnodes representing its legs.

Furthermore, objects may be assigned with so-called materials – fixed properties lists describing how particular objects should be rendered.

A primitive scene graph's rendering algorithm consists of visiting every node of the tree and rendering each of the objects individually using material properties.

Many of the modern graphics engines provide more sophisticated visualization models. VTK (Visualization Toolkit) [2] is one of those. VTK is made to be extendible multi-purpose engine, and it can be used efficiently in many application areas. Except for high-performance interactive visualization. The approach taken by VTK to support universality is not built on top of graphics pipeline principles and may be lacking required optimization. Also, VTK is essentially a framework (and most graphics engines are), meaning that it enforces own data model and programming methodology onto the user, which may not be always desirable.

2. Analysis

The following figure (fig.1) summarizes the structure of the visualization process and differences between two approaches mentioned above.

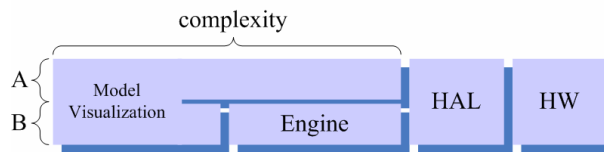


Fig. 1. Visualization process

HW and HAL stand correspondingly for graphics hardware and hardware abstraction layer represented by DirectX or OpenGL. Engine stage is the part of the process taken over by a graphics engine.

The rest of the process is marked as model visualization. This stage is directly connected to the very visualization problem in question. Computer graphics, like any other computing activity, works with **computer models**. These models may represent any kinds of real or imaginary entities, phenomena or processes. Model visualization is essentially a set of algorithms translating the model into another form of representation which is suitable for automatic rendering.

Earlier, we described two common visualization approaches: pure graphics pipeline rendering and graphics engines. They are marked on the scheme as A and B respectively. It is emphasized how graphics engines cut off a serious portion of implementation complexity. However, the real situation is not that simple.

We have mentioned that graphics engines is a whole class of non-standard software. And, moreover, this software is built on the base of certain fixed assumptions, and implements fixed input models that aren't necessarily compatible with the current target model. This happens, for example, in the case when the engines specialize in solid 3D geometric object rendering, and the Model represents a flat user interface.

In this situation software engineers can take one of the three ways:

1. change the model so it fits an engine;
2. provide an engine-model adapter layer;
3. fallback to the pure pipeline rendering.

The first option probably needs the least effort, but the outcome of visualization may seriously differ from the initial expectations as the visualized model gets distorted.

The second option tries to preserve the model's consistency with additional translation stage (fig. 2).

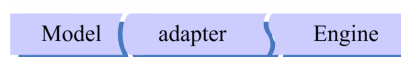


Fig. 2. Model-engine adapter

This may work in some situations depending on how different the model and engine are. If the difference is too big, the adapter itself becomes cumbersome and unmaintainable. In this case the third option becomes preferable: the visualization problem gets solved from the scratch.

The conclusion of this is that we cannot rely on graphics engines from a general perspective. Tasks and models change all the time and engines become obsolete. The variety and quantities grow, which make it troublesome to find the proper match. And so, we have a question: whether anything can be done here in order to improve the visualization process construction.

In our previous works [3, 4] we suggested that it was possible to develop a general methodology for high-level visualization abstractions. That is, to create a model-independent language for visualization process construction.

Together with corresponding support layer libraries, the new visualization process construction would change in the following way (fig. 3).

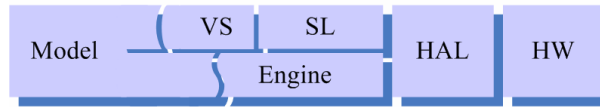


Fig. 3. A new way of the process construction

In this scheme, the model is rendered using an engine with either the first or the second method. SL stands for the standard layer, and VS – for visualization system.

Standard layer is a set of support rendering libraries based on visualization algebra methodology [4]. The library provides tools for process construction in model-agnostic data-driven way. Less effort (in comparison to pure pipeline development) is necessary to make a model adaptation visualization system. And, though the implementation complexity remains slightly bigger than that of the graphics engines, this technique has the advantage of preserving the pipeline's flexibility together with higher level of language abstraction.

3. Object Shaders

Before we start describing the details of the proposed technology, a few words needs to be said regarding how the graphics pipeline is programmed in general.

The graphics pipeline consists of several stages including (DirectX11 model [5]): Input Assembler, Vertex Shader, Hull Shader, Tessellator, Domain Shader, Geometry Shader, Rasterizer, Pixel Shader, Output Merger.

The stages implement different parts of the rasterization algorithm and provide some additional functionality. Shader stages are programmable; all the others are configurable. Configurable stages implement fixed algorithms with some adjustable parameters. Programmable stages, in their turn, can be loaded with custom programs (which are called shaders) implementing specific custom algorithms. This is what essentially gives the pipeline its flexibility.

Shader programs of different types operate with different kinds of objects. Vertex shaders operate with geometry vertices and control geometrical transformation of 3D model and its projection onto 2D screen. Pixel shaders work with fragments – pieces of the resulting image that are later translated into pixels.

The most common way in writing the shader programs is to use one of the high level shading languages: HLSL for DirectX, or GLSL for OpenGL. These languages have common notation and similar instruction sets determined by underlying hardware. By their nature, the languages are vector processing SIMD-based. And corresponding shader programs implement algorithms that describe how every vertex or fragment needs to be treated independently, enabling massive parallel data processing capability.

In our work, we pursue technological integration into the graphics pipeline, rather than its replacement with another set of tools. That is why the technical realization of the proposed methodology is based on emulation of an additional programmable pipeline stage which we call Object Shader.

Object Shader stage is a broad name for three types of programmable components: pipeline, sampling, and rendering. These components are based on corresponding notions from visualization algebra [4]: visual expression, sample, and render operations.

Visual expression in visualization algebra (VA) is a formalized algorithm operating in a generalized object space. Objects in VA are represented with tuples of attributes projected onto a model-specific semantic grid. The methodology does not make any assumptions on the nature of objects and their content, treating all data equally.

Visual expressions in VA are constructed using four basic operations:

1. sample – object subset selection,
2. transform – derivation of new objects on the base of existing ones,
3. render – objects translation into an image,
4. blend – operations with the resulting images.

The final expression for target model visualization must have one input (all the model's data) and one output (the resulting image).

On the technical side, the expression is represented with a program on a language similar to HLSL. The additions are: data type **ptr** used for declaration of resources, object types Scene, Objects and Frame, various rendering-specific functions.

Object layout declaration in object shaders is done in a common way with structures:

```
struct Object
{
    field-type field-name : field-semantic;
    ...
};
```

Sampler functions are simple routines returning boolean values:

```
bool Sample(Object object) { return sampling-condition; }
```

Render procedures generate operation sequences for objects of the supported type and return Frame as a result:

```
Frame Render(Object object)
{
    instruction;
    ...
    return draw-instruction;
}
```

Pipeline procedures combine sampling and rendering operations into the final visualization algorithm. The input of a pipeline procedure is a Scene object and a Frame object is output.

```
Frame Pipeline(Scene scene)
{
    Objects list = Sample(scene);
    ...
    return Render(list);
}
```

Visualization system routes data streams according to the pipeline procedure logic, splitting it with samplers and processing with renderers. The unit routines are designed to be atomic in the same way how it is done for the other shader types: processing one object at a time, allowing massive parallelization.

This technique is fairly similar to effects in DirectX [6]; but, at the same time, the differences are apparent: effects cannot be used for building the complete visualization pipeline.

In the last part of the paper we will go through the real usage example.

4. Usage Example

The sample model consists of one 3D object (a building) and one 2D object (overlay image). The resulting visualization should visualize the building at the center of screen and make it possible to change its orientation. The overlay image should be drawn at the top right corner.

From the model description we know that there are two types of objects and therefore we need two sampling and two rendering procedures.

Sampling procedures:

```
bool Sample1(int type : iType) { return type == 1; }
bool Sample2(int type : iType) { return type == 2; }
```

The procedures get the type field from object stream and compare it against predefined constant values. So the first procedure will sample 3D objects (type 1) and the second – 2D objects (type 2).

Rendering procedure for 3D objects starts with object layout declaration:

```
struct Object
{
    float4x4 transform : f4x4Transform;
    ptr VB : pVB;
    ptr VD : pVD;
    ptr tx0 : pTX0;
    int primitive_count : iPrimitiveCount;
    int vertex_size : iVertexSize;
}
```

```
};
```

The supported objects must have transform matrix, vertex buffer (VB), vertex declaration (VD), texture (tx0) and common geometry information: primitives count and vertex size.

Then, we declare external variables, which are required to be provided by the user: common vertex shader (VS), common pixel shader (PS) and view-projection transformation matrix (f4x4ViewProjection).

The rendering procedure itself:

```
Frame OS_Basic(Object obj)
{
    SetStreamSource(0, obj.VB, 0, obj.vertex_size);
    SetVertexDeclaration(obj.VD);
    SetVertexShader(VS);
    SetPixelShader(PS);
    SetVertexShaderConstantF(0, &obj.transform, 4);
    SetVertexShaderConstantF(4, &f4x4ViewProjection, 4);
    SetTexture(0, obj.tx0);
    return DrawPrimitive(4, 0, obj.primitive_count);
}
```

The procedure makes a number of state change calls and invokes a drawing routine. The second rendering procedure OS_UI is implemented in a similar way.

The resulting pipeline procedure is very simple: it needs to use the declared samplers and renderers, and combine their output:

```
Frame PP_Main(Scene scene) { return OS_Basic(Sample1(scene)) + OS_UI(Sample2(scene)); }
```

And the following picture represents the result of visualization (fig. 4):



Fig. 4. Result of visualization

Conclusion

Computer visualization still holds the status of a heavily evolving scientific and engineering area. Dozens of new techniques and hardware emerge every year. And, with further advancements, this environment may require certain intensive changes in order to stay comprehensible and maintainable.

This paper provides justification and a short overview of the technological implementation of the new visualization methodology based on so-called visualization algebra. This methodology has a potential to improve the most popular existing methods of visualization in terms of flexibility and accessibility.

In difference from graphics engines, the proposed approach is based solely on the graphics pipeline extension. It preserves the basic methodology and does not introduce any new undesirable constructs.

References

1. Gomes, J. Computer Graphics: Theory and Practice / J. Gomes, L. Velho, M. C. Sousa. – Peters/CRC Press, 2012. – 544 p.
2. Visualization Toolkit [Electronic resource]. – Mode of access : www.vtk.org. – Date of access : 5.05.2014.

3. Krasnoproshin, V. Graphics pipeline automation based on visualization algebra / V. Krasnoproshin, D. Mazouka // Proc. of Intern. Conf. on Pattern Recognition and Information Processing. – Minsk, 2011.

4. Krasnoproshin, V. Novel Approach to Dynamic Models Visualization / V. Krasnoproshin, D. Mazouka // Journal of Computational Optimization in Economics and Finance. – 2013. – Vol. 4, Issue 2–3. – P. 113–124.

5. MSDN article [Electronic resource]. – Mode of access : <http://msdn.microsoft.com/en-us/library/windows/desktop/ff476882%28v=vs.85%29.aspx>. – Date of access : 5.05.2014.

6. MSDN article [Electronic resource]. – Mode of access : <http://msdn.microsoft.com/en-us/library/windows/desktop/ff476136%28v=vs.85%29.aspx>. – Date of access : 5.05.2014.

DATA MINING AND KNOWLEDGE- BASED TECHNOLOGY

V. Krasnoproshin¹, V. Obratsov¹, H. Vissia², Q.T. Nguyen³

¹Belarusian State University, Minsk

²Byelex B.V., Rotterdam, The Netherlands

³Ho Chi Minh City University of Education, HCM City, Vietnam

e-mail: krasnoproshin@bsu.by

This paper highlights an approach for building information-computer technologies for solving problems based on knowledge.

Introduction

The development of information technologies associated with computer data processing (IC-technologies) has several stages, from data structures (DS) to databases management systems (DBMS). The range of problems that can be solved has become wider while the problems of standardizing development stages and program maintenance have disappeared.

Today <DB+DBMS> technology is widely-popular among IC-technologies. The rest are either devoid of their integrity and completeness or are of specialized nature in a technical or applied sense. It is quite evident that all these technologies do not fully satisfy the whole range of existing and newly appearing problems.

Development of hardware, host systems, IC-technologies and means of communication has made it possible to involve in automation problems related to intellectual activities of a human being. These problems deal with management, forecast, diagnosis, etc. where it is necessary to process not only data but the knowledge as well. The knowledge as an object of formal description and means for solving problems has long been used in scientific investigation. For example, in artificial intelligence (AI) much experience has been gained for stating and solving problems based on knowledge. In early 1980s a technology was proposed that supported computer-assisted knowledge processing, i.e. expert systems (ES). But as a technology it was not a great success primarily due to the lack of continuity with respect to previous technologies. This is the key to understanding what features should characterize a technology of knowledge processing. In our opinion such a technology should be based on the idea of separation between knowledge and means of its handling. This leads to the necessity of introducing the terms of knowledge base (KB) and systems of their management (KBMS). The paper presents justification of such an approach. It describes a real operating system (solution of a set of medical problems related to orthopedics) that can be considered as a prototype of KBMS.

1. Knowledge and computer technology

What is a technology of knowledge processing? To answer the question we need some additional definitions and refinements.

Let's start with evident notions that can be found in any textbook on data bases, object-oriented programming or artificial intelligence. To make the description concise we will represent them as theses.

1. A necessary condition for the use of computers (computer technologies) when solving problems in any subject area (SA) is a change-over to the formal level. To do this it is necessary to build a conceptual representation of SA in terms of <object, relation>. Here by an object is meant a set of elements having the common structure and the way of building in the following sense:

$$\text{object: } P_1 \times \dots \times P_n \rightarrow D_1 \times \dots \times D_n,$$

where P_i are signs, D_i is a set of possible values of a sign ($i=1, \dots, n$). Relation is a subset of Cartesian powers of the set {object }¹, that can be built for each object.

The construction shown above is a set-theoretic one and it is good as abstraction. In practice, use is made of its various equivalents, where <object, connection> is the most common choice. The result is a construction that is very close in meaning to a semantic network. Thus it is possible to say that the basis for formal description is semantics of the SA.

2. A sufficient condition to use a computer technology is a statement of a problem, i.e. determination of a purpose of object existence. In so doing the purpose itself and means of achieving it substantially depend on semantics and properties of <object, connection> elements (let denote it for brief

<OC>). A purpose is always associated with a certain subset of $\{\langle OC \rangle\}^k$ and is achieved as a result of performing the algorithm

$$A: \{\langle OC \rangle\}^m \rightarrow \{\langle OC \rangle\}^k.$$

The entry and exit of the algorithm are Cartesian powers of <OC> elements.

3. The foregoing relates to information. An IC-technology is a set of means, which provide functioning of certain information. It is evident that a technology should be to a maximum extent invariant with respect to the content of information. That is why it can be determined through the properties of elements of conceptual representation about the SA. From this standpoint let's analyze <DS+PL> and <DB+DBMS> technologies (see the table below).

Table

	<DS+PL>	<DB+DBMS>
Objects	sets of vectors, matrices and all possible structures above them (lists, queues, etc.)	sets of N-tuples
Connections between objects	only connections of order are determined	besides connections of order, connections of type are determined (they are called - 1:1, 1:N, M:N)
Algorithms for obtaining new <OC> samples	built-in algorithms of ordering	besides algorithms of ordering, use is made of algorithms based on connections of type
Means of implementation	PL	DBMS

The following features of the described technologies are worth mentioning:

- object properties are determined indirectly through properties of connections. In this sense the notion of an object for <DB+DBMS> technology is broader, almost corresponding to the notion of a class in OOP [2];

- algorithms are a consequence of agreements about object properties and connections. These algorithms form the kernel of the technology, its indispensable part. It is obvious that ways of obtaining objects, which have not direct relation to the corresponding connection, can be realized by means of PL and DBMS. To accomplish this, DBMS contains a typical built-in PL apart from a standard part (e.g. SQL).

The performed analysis allows to make the following conclusion: between the considered technologies there is a dependence of <DS+PL> < DB+DBMS> type. The extension of its possibilities takes place only at the expense of properties of the connection. The rest is a consequence.

Now we have everything necessary to determine knowledge as a new object of a computer technology. First, let's introduce the notion of data. Taking into consideration the reasoning made above.

By data is meant a part of information about SA for which the following items are determined:

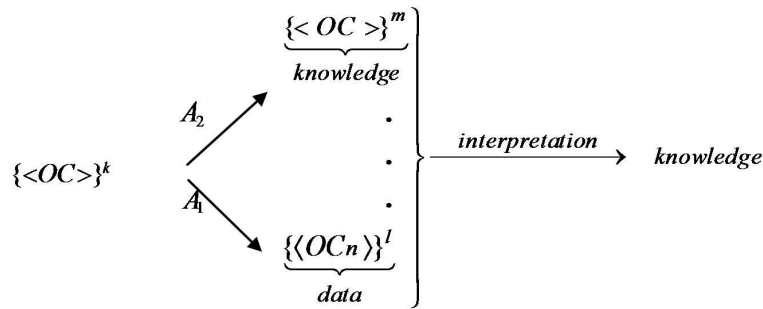
- description at the conceptual level of the structure $\{\langle OC \rangle\}^t$ (t is a random, possibly, an infinite number);

- a connection that is not more complex than a connection of type;

- algorithms conditioned by such connections.

Knowledge also has an algorithmic origin because it is a basis and a result of transforming one part of $\{\langle OC \rangle\}^k$ structure into another one. Schematically this can be shown in the manner below.

The reasoning made above allows to determine specific character of knowledge at the level of $\{\langle OC \rangle\}^k$ structure as well as at the level of algorithms. For this purpose it is necessary to extend the notion of connection in data definition. If maximum generalization is taken into account, then this can be transformation of a connection into a connection with all properties of objects. It correlates with the ideology of OOP. But in problems of knowledge processing the connection itself can be an aim. And finally, only in this case one can expect that the knowledge handling technology will contain components of previous technologies.



where A_1 are algorithms mentioned in data definition, A_2 are algorithms for obtaining knowledge.

It is an easy matter to determine also the character of such a connection. It will suffice to recall that at the highest level of abstraction of any set-theoretic structure there is logic, i.e. is a connection should specify logical dependencies between parts of $\{\langle OC \rangle\}^k$ structure. As this takes place, mathematical tools for analyzing the structure and building algorithms should generalize the relational algebra. These can be algebra systems that contain a language allowing to describe logic elements and build algorithms.

Now it is possible to introduce the notion of knowledge.

By knowledge is meant a part of information about SA for which the following items are determined:

- description at the conceptual level of the structure $\{\langle OC \rangle\}^l$;
- a connection that is an object by itself;
- algorithms conditioned by connections.

The connection should establish logical dependencies in $\{\langle OC \rangle\}^l$ structure and, as an object, should be an aim for solving the problem. The attainment of the aim (solution of the problem) should be supported by algorithms of deduction.

Now to make knowledge an object of a technology it remains to determine the notion of a knowledge base.

By KB is meant a structure $\{\langle OC \rangle\}^k$, in which a connection can be of any kind mentioned above (of order, type or a logical one) in any combination and with clearly specified aims for solving problems (algorithms). KBMS, in its turn, is an environment for supporting such a structure with possibilities of interpretation.

2. Main problems of <KB+KBMS> technology

Now let's consider problems that arise when developing a technology. As in previous cases the technologies can be divided into two interconnected and interrelated groups.

1. Problems at the level of KB.

There are problems of representation and algorithmic processing of knowledge (data). These problems are interconnected because the nature of algorithmic processing in all cases is specified by a connection.

The problem of knowledge representation is considered at three levels: conceptual, logical, and physical. This results in solving a problem of designing a particular KB for its subsequent storage, adaptation and knowledge (data) handling. Let's discuss each of the levels in more detail.

At the physical level data and knowledge are not differentiated, because in all cases they deal with memory cells. That is why here it is sufficient to have mechanisms (files, keys, indexing, etc.) that have been realized in previous technologies. The conceptual level may be based on mechanisms developed in DB and OOP. Most likely this will be a logical structure built on the basis of a universal object (relation) [1]. For formalization it is possible to use the language of algebraic systems that contains as subsets languages for describing functional and logical dependencies. An element of such structure can be a class [2].

The logical level of knowledge (data) representation is at base of the whole KB concept. It is necessary to build a model of KB, i.e. to design a logical structure of classes and objects, to determine processes of their interaction [2], as well as a logical schema of data [1]. To develop a logical schema of data use is made of standard procedures of normalization. As for the methodology of designing a logical structure of classes and objects, it has not so far the necessary universal mechanism and it is not standardized. Standardization is proposed to carry out by introducing a new restriction: completeness, independence and closure of knowledge are introduced in addition to data integrity. For standardization an important point is that a connection (understood as in OOP) at the level of knowledge is at the same time

an object at the level of data. Thus the object can be subjected to normalization and handling typical for DB.

The problem of algorithmic processing of knowledge (data) is based on classification of connections. In our case it is necessary to specify the nature of algorithmic processing only for connections that are not supported in a standard manner in OOP. These connections relate to logical dependencies or more precisely to inference. For the chosen model of knowledge static and dynamic aspects of such inference are determined at the level of designing a physical structure by refining the architecture of processes [2]. With respect to content the inference can be deductive or inductive (the latter one is important due to incompleteness of knowledge typical for actual practice, empirical nature of knowledge, etc.). Both the ways of inference should be supported by KBMS means. Deductive inference, the corresponding procedures for knowledge representation, resolution (as an algorithm of inference), requirements to knowledge and other questions have been studied in sufficient detail. It is clear that without inductive inference our technology will be incomplete. Because of this, the problem of developing procedures for inductive inference and essentially the unification of inference becomes the key one. It is closely related to such problems as: choice of a language, a procedure for describing knowledge, relocation and functioning of inductive resolution [3].

2. Problems at the level of KBMS.

At this level it is possible to single out two interrelated groups of problems dealing with design support and KB functioning. There are also important and complex problems of KBMS technical realization (architecture of modules, a programming system, etc.). But taking into account the methodological aspect of the paper we will consider only expected functions and some general requirements to KBMS.

A good prototype of KBMS can be any DBMS in the sense that functions of the latter one can be extended to design support and KB functioning. In general terms, this can be done in the following directions.

At the stage of design:

- a) by introducing and supporting classes and their particular samples (objects);
- b) by introducing and supporting additional types of connections, typical for OOP, along with logical dependencies (as new objects);
- c) by extending the algorithmic base for analysis and handling of KB. This should result in construction of a corresponding extension of a language of SQL type (or a similar one);
- d) by forming an interface part of KB (as dictionaries of SA, etc.) to provide subsequent interaction of the user with the finished system.

At the stage of functioning:

- a) a possibility of KB adaptation to varying data and set-up to new knowledge (additional learning, self-learning);
- b) operation in different modes selected by the user;
- c) a possibility to control user's actions and to explain any step of system's operation.

There is no doubt that the above list of functions is only schematic. At present none of the systems [3], to which a reference can be made, has the above features of KBMS. In this connection it is also of interest to consider the relationship between KB models, that will be supported by a particular KBMS, and existing ones (e.g. frames, semantic networks, etc.). Such and similar questions should be solved when developing KBMS.

Taking into consideration the above, one would expect that relationship between technologies will be continued, i.e. $\langle \text{DB+DBMS} \rangle \subset \langle \text{KB+KBMS} \rangle$. And, consequently, a wider range of practical problems will be solved by means of the latter one.

3. An example of a system based on knowledge

Let's consider solution of a practical problem for which "ORTHO-EXPERT" (OE) system has been developed and which in functional sense is oriented to complete information provision of clinical practice of an orthopedist. It is just in the process of working on the project that questions have arisen which have been described above. The obtained solution is an integral one, but in this paper we consider only the part that bears a direct relation to $\langle \text{KB+KBMS} \rangle$ technology. Additional information one can find in [3].

The system consists of two universal applications - Ortho-Expert and Ortho-Expert Designer, developed according to a common technology within one project. Each application solves a specific range of problems. In particular it ensures:

- operation of the system with various information, i.e. knowledge about a subject area, metaknowledge, references in different representation (text, graphical, audio and video). As this takes place, the information should be semantically ordered depending on the context of its usage;
- flexibility of the system and simplicity of its adaptation for processing new types of data and knowledge;
- portability when using different platforms for the system (now the system operates in Windows 7 environment).

The technology of system construction is based on the principles of object-oriented programming. Knowledge handling is carried out in the system at the stages of input (Ortho-Expert Designer) and manipulation (Ortho-Expert).

Conclusions

Taking into consideration the above it is possible to make a conclusion that OE is a good prototype of KBMS. If it is granted that KBMS is really a next step in the development of the computer technology and general features of the latter are described correctly in the paper, then the legitimacy of the conclusion is beyond question.

We can argue for the made conclusion in a simple manner. The main purpose of <KB+KBMS> computer technology is in separation of knowledge from means of its handling. But this is possible only when specifying algorithms of knowledge processing and in fact it has been realized in the process of building OE.

At present OE does not fully correspond to characteristics of a system of KBMS type. Two directions of further work are possible:

- to add possibilities of DBMS to the existing variant of OE;
- to add possibilities of OE to the existing variant of a DBMS.

The most practicable is the first one. At least there are all prerequisites for its realization: there should be decisions at the level of state government.

References

1. Maier, D. *The Theory of Relational Databases* / D. Maier. – Computer Science Press, 1982.
2. Booch, G. *Object Oriented Design. With Applications*. The Behjamine / G. Booch. – Cummings Publishing Company, Inc., 1991.
3. Krasnoprosin, V.V. *Problem of Solvability and Choice of Algorithms for Decision Making by Precedence* / V.V. Krasnoprosin, V.A. Obraztsov // *Pattern Recognition and Image Analysis*. –2006. – Vol. 16, no 2. – P.155–169.

STANDARTIZATION OF MOBILE HETEROGENIOUS OBJECTS MONITORING BASED ON PATTERN RECOGNITION

A.I. Kuzmich

Belarusian State University of Informatics and Radioelectronics, Minsk

e-mail: kai57@list.ru

The problem of mobile heterogeneous objects unification description is considered. As a solution to the problem, we propose to use object's graph model and visual standards for her elements. The mechanism for constructing model and standards are described.

Introduction

Mobile heterogeneous object (MHO) is the material or virtual structure, which consists of heterogeneous components and changes its location in a natural or artificial environment [1]. The report examines the material MHO. An example of such objects is a train, which components are the locomotive, cargo beds and tanks.

Successful implementation of business projects involving MHO depends largely on an effective monitoring. Monitoring (in the broadest sense of the word) means a system for assessing the current state of the MHO and synthesis of the corresponding management solutions [2]. Despite the abundance of monitoring techniques and systems (mainly for individual components), a number of pressing problems is not enough studied. The major ones include the following:

- no typical formula describing MHO clear to managers at all levels of the project exists;
- a mathematical model of the MHO, “clear”, both to a manager and a computer system monitoring is not developed;
- standards to visualize the results of monitoring, invariant to project specificity and intuitively understood by the manager have not been developed;
- a typical monitoring system architecture based on standards haven't been developed.

The report discusses the solutions to these problems based on the theory of pattern recognition.

1. Formulation of the problem

Let there be a company implementing business processes using different MHO types.

It is required:

- to develop a universal formula describing MHO structure;
- to build a MHO model that provides a mathematical processing;
- to develop standards for the visualization of model elements in a form intuitively understood by managers at all levels of the project implementation;
- to develop monitoring system architecture based on models and standards.

Basic requirements for the solution are: description of the MHO should be invariant to the type and functionality of the facility; MHO model should provide the implementation in modern programming languages (C #, Java, C + +) and operate results on different types of devices (PCs, tablets, smartphones).

As a general approach, we use the theory of pattern recognition.

2. MHO description formula

MHO components feature extremely large variety. Nevertheless, according to the basic paradigm of synergy, for many objects and processes different in nature a universal description can be found [3]. In the case of MHO common managers interest at all levels in building an intuitive unified MHO model meeting the requirements of the project and providing monitoring tasks solutions automation, including visualization is obvious. To standardize the description INR offered an option that can be represented as a tuple:

$$\text{MHO} = (\text{name}, N, \langle I, P, Q \rangle), \quad (1)$$

where: name is MHO name; N is the total number of components; I is the serial number of the component; P is the type of component; Q is the number of components of this type. This option provides a simple description of the MHO on any modern programming language.

As an example of the formula (1) usage, let's describe the MHO named RVP1.4 in C#, comprising eight components of three different types: 1 locomotive, 4 tanks, 3 cargo platforms:

string MGO = “RVP1.4, 8 (1, locomotive, 1; 2, tanks, 4; 3, cargo platform, 3)”.

For practical use of the MHO formula it is necessary to display it in a mathematical model that is convenient for implementation on a computer.

3. MHO graph-model

As a mathematical model of the MHO we will use the graph as it is an universal form of representation hierarchies [4].

Graph will be formed on the parsing basis (analysis and decomposition) of the formula (1):

Step 1. Perform formula parsing to allocate elements: name, N, I, P, Q.

Step 2. Form the vertex of the graph (name).

Step 3. Determine the amount and form nodes of the graph (P).

Step 4. Determine the number of nodes and to form terminal nodes (Q).

As a result of this algorithm with a string of MHO, we obtain the graph (fig. 1).

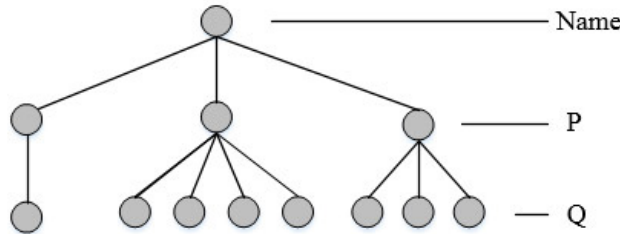


Fig. 1. MHO structure

It is obvious that this option MHO representation is quite convenient for implementation on a computer, but is hard to understand by decision-maker, especially in case of their rotation. To provide a thorough understanding it is necessary to create a series of simple corporate standards.

4. Components images unification

The greatest interests for the manager are the components that are involved in the MHO and are visually well known to him. So first of all let's solve the problem of understanding. For this we introduce a corporate standard representation of the MHO components types, including type identifier, its name and the file name with its image. As an example, the corporate standard description of the components types for railway depot is given below (table 1).

Table 1

Components visualization standard

Component type	Component type identifier	File name
T1	Locomotive	Loc.jpg
T2	Boxcar	Vt.jpg
T3	Passenger car (κ)	Vk.jpg
T4	Passenger car (π)	Vp.jpg
T5	Cargo platform	Pg.jpg
T6	Oil tank	Cn.jpg

This standard is saved in a corporate database, it is easy to edit and adapt to the specifics of any company. It's enough to change the component type identifier and file names with graphic images.

As a result of this standard implementation, MHO description string is significantly reduced:

MHO string = “ЖДЦ11.4, 8 (1, T1, 1; 2, T6, 4; 3, T5, 3)”;

After parsing and reading this string from a corporate graphics components images database, we obtain more informative (compared to fig. 1) version of the MHO structure (fig. 2).

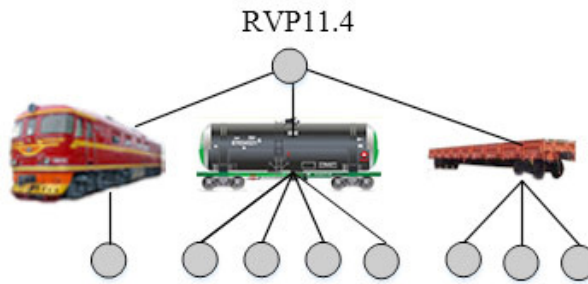


Fig. 2. MHO partial visualization

This visualization option has the obvious drawback – the manager cannot visually determine the current MHO status and its components. To eliminate this, we introduce another corporate standard, unifying states visual representation.

5. States images unification

The number of MHO states, as well as any physical object can be infinitely large, but the sphere of interests of decision-makers includes only those states that may affect the possibility of the project implementation. Analysis of real projects using the MHO implemented by cargo-carrier companies on the Latvian Railway showed that in most cases only three or five states are enough for the decision-maker. These figures are in good agreement with the results of psychological research, described in [5]. Taking the maximum value as a basis, we introduce the standard representations of five states; represented as a variable of (string) type and graphic icons (jpg). To describe these states we will apply states gradation clear for everyone: from “excellent” to “destruction”, the corresponding constant (V), and a numerical criterion (K). To construct icons it is proposed to use common “allow”, “warning” and “prohibit” colors – green, yellow and red. To construct the border states we will divide icons in color in two parts by the principle “from good to break”. The construction result of the relevant standard is presented below (table 2).

Table 2

States visualization standard

Type	State numerical criterion	State value	State identifier	File name	Color identifier
V1	$K \approx 0.00$	“excellent”	excellent	G1.jpg	Green
V2	$K \approx 0,25$	“good”	good	G2.jpg	GrYel
V3	$K \approx 0,50$	“average”	average	G3.jpg	Yellow
V4	$K \approx 0,75$	“break”	break	G4.jpg	YelRed
V5	$K \approx 1.00$	“destruction”	destruction	G5.jpg	Red

Standard elements can be used in MHO control tasks, including the organization of parametric dialogue [2]. If necessary, the standard is easily adjusted and adapted to the specifics of any company. To do this it's enough to change the state values or file names with icons images.

As an example, we use this standard to visualize the MHO current state, obtained by K coefficient random selection and obtain the following result (fig. 3).

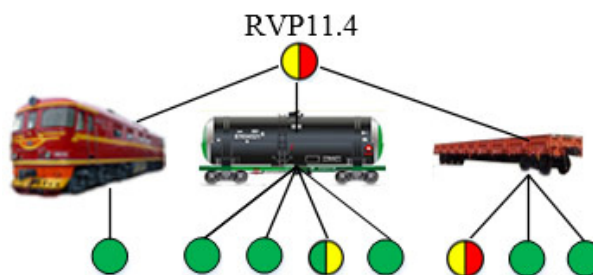


Fig. 3. MHO complete visualization

Scheme gives the manager a complete picture of the MHO: first cargo platform is in pre-emergency state. State of the whole train depends on the state of its weakest element, so MHO state is generally rated as “break”, i.e. pre-alarm. By implementing this approach in the online version, the project manager can

monitor the current state of the MHO through any devices with the ability to connect the net (tablets, smartphones, etc.).

As a result of this standard application, visualization result becomes more informative and allows the decision-maker to understand the condition of the MHO and its components in a blink of an eye. Accordingly, the reaction time of the decision-maker to the situation and synthesis of management solutions is reduced.

6. Monitoring system architecture unification

Unification of monitoring systems architecture will be carried out in two stages. The first stage is the standardization of monitoring tasks based on pattern recognition theory. The second is the construction of a model based on multi-agent architecture approach.

In accordance with the standard object recognition task [6], the task of monitoring represent in a form of tuple:

$$\text{Task} = (X, f1, f2, E, f3, V, f4, U, f5, V, U), \quad (2)$$

where X is the object diagnostic characteristics; $f1$ is an algorithm for X acquire; $f2$ is an algorithm for constructing etalons; E – etalons; $f3$ – an algorithm for constructing states; V – states; $f4$ – an algorithm for constructing management solutions; U – management solutions; $f5$ – recognition algorithm state of the object; V, U – current status and management.

In accordance with the multi-agent approach [7], monitoring system architecture, including the tuple (2), will consist of five software agents:

$$A1 (f1, X) \rightarrow A2 (f2, E) \rightarrow A3 (f3, V) \rightarrow A4 (f4, U) \rightarrow A5 (f5, V, U).$$

This version of the architecture is of a hard character, but provides a flexible implementation and improvement of the code of each agent without changing the other agents.

Conclusion

The report presents the following results:

- MHO compact description formula, providing construct a basis for automation of monitoring tasks is proposed;
- MHO mathematical model corresponding to the formula is developed;
- standards for describing the elements of the MHO component types and their states, provides an intuitive visualization of the state of the MHO as a whole and each of the components is proposed;
- an example of application of the formula and the standards for typical for rail transport MHO description is given;
- unified variant of architecture monitoring system that combines corporate standards and standards for constructing agents is proposed.

References

1. Рушкевич, А. Мониторинг подвижных объектов / А. Рушкевич, В. Осадчий // Беспроводные технологии. – 2010. – № 3. – С. 56–60.
2. Kuzmich, A.I. Remote monitoring system for mobile objects / A.I. Kuzmich, G. Shakah, A.N. Valvachev // Proc. of 10th Intern. Conference (PRIP'2011), Minsk, 2011. – P. 427–430.
3. Чернавский, Д.С. Синергетика и информация. Динамическая теория информации / Д.С. Чернавский. – М. : Едиториал УРСС, 2004. – 288 с.
4. Бурков, В.Н. Теория графов в управлении организационными системами / В.Н. Бурков, А.Ю. Заложнев, Д.А. Новиков. – М. : Синтег, 2001. – 124 с.
5. Кобринский, Б.А. К вопросу о формальном отражении образного мышления и интуиции специалиста в слабоструктурированной предметной области / Б.А. Кобринский // Новости искусственного интеллекта. – 1998. – № 3. – С. 64–76.
6. Murty, M. Pattern Recognition / M. Murty. – Springer, 2011. – 275 p.
7. Weiss, G. Multiagent systems / G. Weiss. – MIT Press, 2013. – 920 p.

APPLICATION OF BIG DATA TECHNOLOGIES TO IMAGE PROCESSING

S. Lazebnikov, A. Boudnik, A. Sarychev
EPAM Systems, USA

Big Data technologies capabilities are introduced and potential areas of their application for Image Processing and Pattern recognition are considered.

When size becomes a problem

The set of technologies associated with Big Data is growing very fast. There are also many interesting research papers in the field of image analysis and pattern recognition. This article by no means intended to provide a comprehensive picture of interaction between these areas of research. The authors have much more modest aspirations – to introduce some of Big Data technologies capabilities and point out the potential areas of their application for Image Processing and Pattern recognition.

From that viewpoint we can identify three phases in the information processing: Algorithmic, Technological and Analytical.

The Algorithmic phase or Image Processing or Pattern Recognition is a process of extraction of so-called features from one object of processing (be it an image or not). Features are numerical representation of, as example, encoded color and texture properties of the image, the layout and geometric shape characteristics of an objects. For example, the photometric and geometric features. Photometric features exploit color and texture cues derived directly from raw pixel intensities. Geometric features are cues such as edges, contours, joints, polylines, and polygonal regions. A suitable shape representation extracted from the pixel intensity information by region-of interest detection, segmentation, and grouping. The algorithms like DHOG, direction of gradient, SURF, SIFT and many others are used at this stage. As at the time of writing this article we do not see how Big Data technologies can be used in this phase. Granted those are very computationally intensive algorithms, but seems like modern computers are keeping up and everything that should be done with regard to processing of one object can be done on one computer.

When there are many objects to be processed and the size of the data becomes part of the problem – this would be the time to call the Big Data. After all they call some of the data Big for the good reason. The Technological phase includes a set of activities dealing with capturing all available objects, invoking object processing code, storing, searching, and retrieving the objects. When only few images need to be processed, all of these activities are relatively trivial. However, as the number of objects to be processed grows the Technological aspect became more and more important. As Roger Magoulas observed: “Data become Big Data when its size becomes a part of the problem”. For example, “...It is estimated, that by 2015, the average hospital will generate 665TB of data... Medical image archives are increasing by 20–40% annually... Today, 80% of data is unstructured, such as images, video and email” [1]. Dealing with challenges of storing, processing, and searching at this scale calls for new architectures and technologies.

The Analysis phase is dealing with the data after it is ingested and stored. For example, healthcare application like PACS (Picture Archival & Communication System), is designed to analyze and provide access to very large volumes of different kinds of medical images. Just ingesting and storing all kinds of data is never the goal. Although one of the funny definitions of Big Data states that the data becomes Big, when it is cheaper to store it than to decide on which part of it can be discarded. Analysis is, obviously, the most important phase. At this stage usually more than one object at a time needs to be processed, interactions between objects are studied and that also requires new approaches to computational models. This is also an area, where the Big Data technologies can prove to be useful.

How the Big Data Technologies can help?

Storage

It would be reasonable to assume that objects, being processed, need to be stored somewhere at some point for the future use. Big Data storage offer inexpensive, fault-tolerant distributed storage, where data is stored across the several (or several thousand) of computers working together. One of key points that allows Big Data storage to operate efficiently is “write-once” principle of operation, meaning that any file is immutable and once it had been written and properly closed, the file cannot be modified. The only way to update the content would be to write updated content in a new file and remove the old one. This approach avoids the necessity of locks and other kinds of inter-process synchronization and this is one of the reason that the data storage is cheap.

Data could be stored locally or remotely on a public cloud.

For the local data center, the most affordable option would be HDFS, Hadoop’s [2] file system that provides very simple interface, similar to what UNIX or Windows file system does. HDFS advantages are:

- very cheap in terms of price per byte: it allows to use commodity computers with directly attached, inexpensive hard disks w/o RAID controllers combined with zero license & support fees;
- near unlimited scalability: it can provide more and more capacity by plugging in more and more boxes w/o stopping the service. Single NAS sometime can work better than small Hadoop cluster, but as data volume grows, so the cost and bandwidth requirements;
- very high network capacity: using two embedded Ethernet adapters HDFS can provide 2GBs per second per computer, which for large clusters gives total bandwidth of more than a terabyte per second at MapReduce layer. It also can serve multiple clients at the speed of its wire;
- very solid data reliability: HDFS focused on ability to cope with hard drives, computers, and network failures. It will maintain enough redundant replicas of the data, so there is no need to worry about hardware crashes. No expensive hardware solutions like RAID controllers, servers with dual power supplies and fault tolerant network equipment. Instead, cheap boxes with cheap SATA disks, cheap NICs and commodity switches – the software will handle the rest!

As for the cloud, there are two options: either to store data in HDFS deployed on cloud computer service like EC2 [3] or to store data on cloud storage services, such as S3 [4] or Azure [5]. The use of computer service may cost extra money, but gives the same file system like interface and allows using computer power for processing, while with storage service the additional processing power is required. For example, images could be stored on S3, but then need to be processed on EC2 using capacity-on-demand approach.

Usually it is not enough (or practical) to just store raw objects. In order to organize effective access to the raw images and do many kinds of processing it makes sense to store the extracted on the first “algorithmic” phase features of the individual objects. The image features data by itself are multidimensional and complex. In addition, different sets of image features could be added during some of the subsequent processing phases. The techniques of sparse tables, supported by NoSQL technologies seems suitable for the purpose. In essence, those technologies operate on the key-value pairs metaphor. Keys are indexed to enable fast search by key. The multiple values could be stored alongside with the image identifier used as a key. It is beneficial to use additional indices to enable search using different dimensions of the object.

Apache HBase is a very rudimentary example of NoSQL technology, which does not support secondary indices, but is sufficient for simple retrieval. HBase could be co-located with Hadoop cluster, stores key-value sets right in HDFS. Other, more sophisticated NoSQL technologies like Cassandra [6] or Accumulo [7] support secondary indices, but require dedicated (not HDFS) file space to store key-values pairs.

Processing

When images appear in real time, either as single image or in a time series as in video, and immediate reaction is required, processing could benefit from the use of Streaming technologies. When immediate reaction is not an essential requirement, then processing would probably, be better performed in the Batch mode.

Batch Processing

MapReduce may sound familiar to some professionals in image processing. It is a key algorithm used to distribute work around a cluster. The original concepts, calculation paradigm, and MapReduce framework are described in [8]. There is a very brief and simplified explanation of how MapReduce works:

The Map. A map transformation transforms an input key/value to an output key/value:
`map(key1,value) -> list<key2,value2>`

- A map transformation returns a list containing zero or more key/value pairs:
- The output can be a different key from the input
- The output can have multiple entries with the same key

The Reduce. A reduce transformation takes all values for a specific key, and generates a new list of the reduced output.

`reduce(key2, list<value2>) -> list<value3>`

The MapReduce Engine. “The key aspect of the MapReduce algorithm is that if every Map and Reduce is independent of all other ongoing Maps and Reduces, then the operation can be run in parallel on different keys and lists of data. On a large cluster of machines, you can go one step further, and run the Map operations on servers where the data lives. Rather than copy the data over the network to the program, you push out the program to the machines. The output list can then be saved to the distributed file system, and the reducers run to merge the results. Again, it may be possible to run these in parallel, each reducing different keys.” [9].

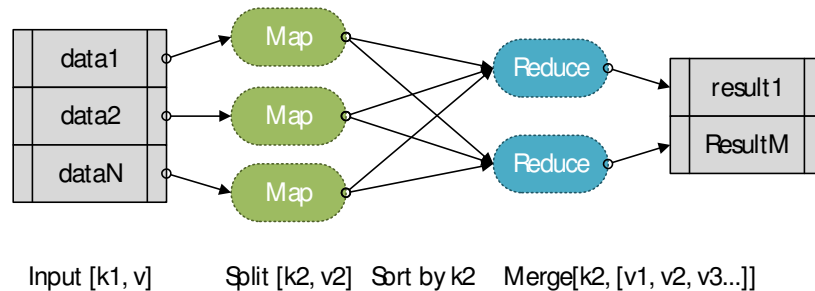


Fig. 1. MapReduce Dataflow

In a degenerate case sort and reduce case may be omitted. The Map-only jobs could be very efficiently used for image feature extraction of image transformation.

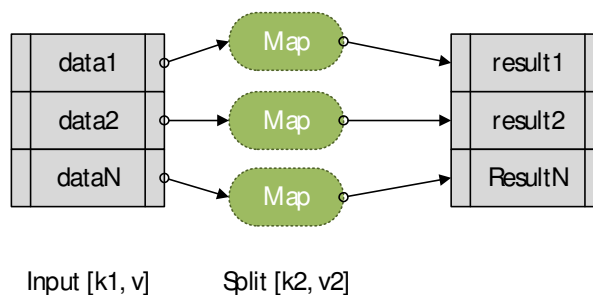


Fig. 2. Map-Only MapReduce Dataflow Image Processing on Hadoop

Hadoop installation could be used to process images out-of-the-box. As an example The New York Times used 100 Amazon EC2 instances and a Hadoop application to process 4 TB of raw image TIFF data (stored in S3) into 11 million finished PDFs in the time slot of 24 hours at a computation cost of about \$240 (not including bandwidth) [10].

Another good example is Skybox Imaging – the using Hadoop for storing and analyzing satellite data [11].

The frameworks like Cascading [12] may enable professionals in image processing to develop Hadoop applications easily by hiding unnecessary details and providing clean Java API.

A specialized framework – HIPI – stands for Hadoop Image Processing Interface is a practical example of how to use Hadoop for image processing tasks in a distributed computing environment [13,14]. There are few challenges this implementation addressed:

- Number of files to store individual images can easily exhaust HDFS capacity, so it uses image bundles (HIB) instead. A HIB is a file, which contains a set of images along with metadata describing their layout. A HIB can be created from an already existing set of images or from stream
- HIPI allows a user to select the images that meet a specified set of criteria. This additional processing step called Cull allows improving the efficiency of some jobs.

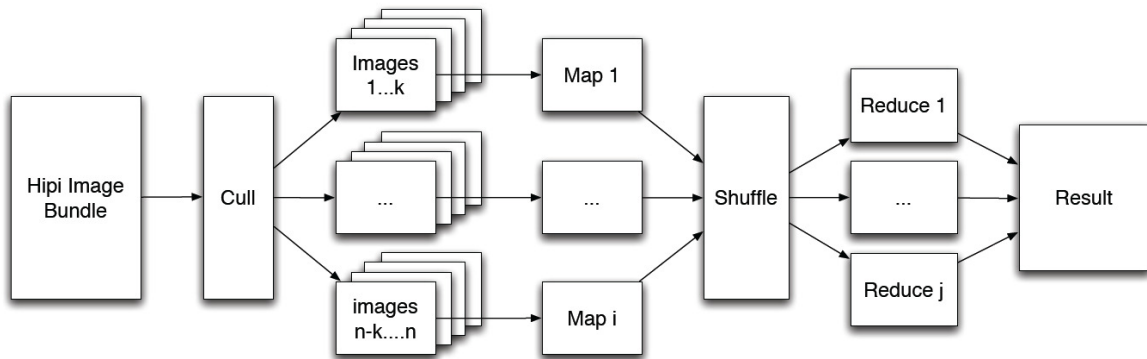


Fig. 3. HIPI Dataflow

1) Hardware acceleration for MapReduce

Nowadays, specialized graphical hardware is widely used to speed-up calculation-intensive tasks, including image processing. The CUDA, stands for Compute Unified Device Architecture improves Hadoop overall performance. For example, mappers and reducers may send data to on-board GPUs. “From the parallel programming point of view, CUDA can help us to parallelize program in the second level if we regard the MapReduce framework as the first level parallelization.” [15]. This implementation uses JCuda – java bindings for CUDA [16].

This approach is beneficent for math-intensive task and useful for image processing, allows to increase performance and reduce power consumption [17].

Another one worthwhile to be mentioned is Mars – a MapReduce Framework on Graphics Processors, where MapReduce paradigm used to organize parallel processing [18].

It might be interesting to implement MapReduce using Hadoop and Mars with uniform parallelism where MapReduce engine will schedule execution of mappers, sorting, and execution of reducers using native GPU threads.

2) Streaming

The Streaming computational model is very simple: extract small input stream “batches” from input stream and minimize overhead per batch to make practical to process small batches. Most of Big Data streaming frameworks keep data in memory to satisfy low-latency requirements, for the obvious reasons it limits the batch size too. In order to cope with peak loads and to provide fault tolerance, streaming frameworks use intermediate messaging buffers like Kafka for to “replay” stream not-yet-processed portion of the stream.

Among the popular stream processing frameworks Spark Streaming [19] is better suited to process image streams because of its ability of state full processing, fast recovery and high throughput.

The key idea behind Spark Streaming is to treat streaming computations as a series of deterministic batch computations on short time intervals. The input data received during each interval is stored reliably across the cluster to form an input dataset for that interval. Once the time interval completes, this dataset is processed via deterministic parallel operations to produce new datasets representing program outputs or intermediate state. Thus, it avoids replication by using lineage for fault recovery. Spark cluster of 100 nodes can process 6 GB/sec of data at sub-second latency (up to 70 MB per second per node).

An implementation of any image processing algorithm could be easily “wrapped” in accordance with Spark framework and run on cluster, without taking care of work distribution, communication and fault recovery. It seems very natural to use CUDA on Spark nodes to speed-up video processing.

Search and Retrieval

The challenges with the image data are not only associated with the size of the image data itself, but also that it is highly dimensional and complex. In order to efficiently retrieve needed object it makes sense to use object features (metadata), which usually are stored in NoSQL database in addition to raw objects. A simple object extraction can search through secondary indices, and then retrieve the raw objects found.

One of the approaches to complex image searches uses notion of image similarity. Assessment of similarities between image features based on mathematical analyses, which compares descriptors across different images. Vector affinity measures such as Euclidean distance, Mahalanobis distance, KL divergence, Earth Mover's distance are amongst the widely used ones.

Query by Image retrieval is a classic example of a search task, when set of features extracted from a given image are used to fuzzy search in image features database using one or more abovementioned similarity metrics.

Case-based retrieval is a more complex task; it is closer to the clinical workflow. A case description, with patient demographics, limited symptoms and test results including imaging studies, is provided (but not the final diagnosis). The goal is to retrieve cases including images that might best suit the provided case description.

Both image-based and case-based retrieval could use machine learning. Mahout [20] could be used as scalable machine learning framework for classification, regression, and clustering.

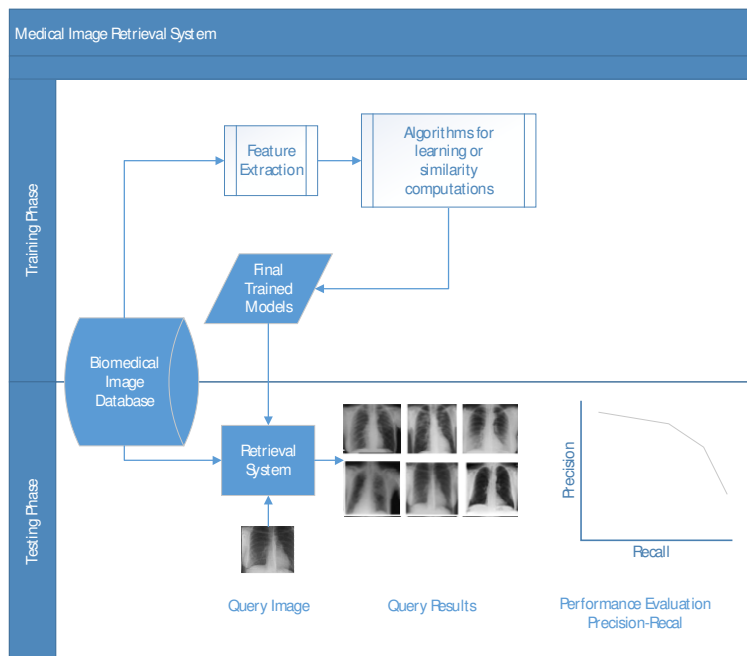


Fig. 4. Biomedical Image Retrieval System

Middleware

It is a very important with a high volume of streaming images to reliably store them for further processing. A Kafka – a high-throughput distributed messaging system provides publish-subscribe service for Big Data ecosystem [21]. Kafka offers strong durability and fault-tolerance guarantees. Brokers in Kafka cluster can handle terabytes of data at speed of hundreds megabytes per second. Its performance grows linear. It could perform simple transformation like format conversion on a fly. It integrates well with other Big Data technologies, for example, can read/write from/to HDFS. The streaming systems like Storm or Spark could be used for further processing after Kafka since they can consume from Kafka and manage the processes doing the consuming.

Data Science and Analytics

There is a difference between capturing images into a Big Data store and getting value of their incredible potential. The Big Data's value can be utilized by intelligently presenting it in relevant and contextual to the problem way.

As example, Skybox customers can then embed their own algorithms in the company's platform and use the analytics engine to crunch the data for their own purposes. Agriculture clients can monitor crop yields. Shipping and supply chain companies can monitor vehicles. Oil and gas companies can evaluate land areas. All of this is based on looking for changes in imaging data by using Hadoop for analytics [11].

As for biomedical imaging, besides storing, searching and retrieving the images, efficiently handling large volumes of medical imaging data and extracting potentially useful information and biomarkers could be beneficial to:

- improve early detection, diagnosis, and treatment;
- predict patient prognosis; aggregated data are used to spot early warning signs and mobilize resources to proactively address care;
- increase interoperability and interconnectivity of health care (i.e., health information exchanges);
- enhance patient care via mobile health, telemedicine, and self-tracking or home devices.

Conclusion

Big Data gives application a freedom of scaling with virtually unlimited space and computing horsepower. The value of Big Data technologies is in their ability to distribute a very complex image processing algorithms to the computer clusters and at the same time to avoid complexity of distributed calculations. The developers can concentrate on application since they are not required to take care of all issues associated with tasks distribution, inter-process communication, synchronization over the network and fault tolerance.

References

1. Healthcare Infographic: The Body as a Source of Big Data [Electronic resource]. – Mode of access : <https://communities.netapp.com/docs/DOC-23102>. – Date of access : 08.05.2014.
2. Welcome to Apache™ Hadoop®! // The Apache Software Foundation [Electronic resource]. – Mode of access : <http://hadoop.apache.org/>. – Date of access : 08.05.2014.
3. Amazon EC2 // Amazon Web Services, Inc. [Electronic resource]. – Mode of access : <http://aws.amazon.com/ec2/>. – Date of access : 08.05.2014.
4. Amazon Simple Storage Service (Amazon S3) // Amazon Web Services, Inc. [Electronic resource]. – Mode of access : <http://aws.amazon.com/s3/>. – Date of access : 08.05.2014.
5. Azure - Discover Microsoft® Windows Azure // Microsoft [Electronic resource]. – Mode of access : www.windowsazure.com/. – Date of access : 08.05.2014.
6. Welcome to Apache Cassandra // The Apache Software Foundation [Electronic resource]. – Mode of access : <http://cassandra.apache.org/>. – Date of access : 08.05.2014.
7. APACHE ACCUMULO // The Apache Software Foundation [Electronic resource]. – Mode of access : <https://accumulo.apache.org/> – Date of access : 08.05.2014.
8. Dean, J. MapReduce: Simplified Data Processing on Large Clusters / J. Dean, S. Ghemawat [Electronic resource]. – Mode of access : https://www.usenix.org/legacy/publications/library/proceedings/osdi04/tech/full_papers/dean/dean_html/. – Date of access : 08.05.2014.
9. MapReduce // The Apache Software Foundation [Electronic resource]. – Mode of access : <http://wiki.apache.org/hadoop/MapReduce>. – Date of access : 08.05.2014.
10. Gottfrid, D. Self-service, Prorated Super Computing Fun! / D. Gottfrid [Electronic resource]. – Mode of access : <http://open.blogs.nytimes.com/2007/11/01/self-service-prorated-super-computing-fun/?scp=1&sq=self%20service%20prorated&st=cse>. – Date of access : 08.05.2014.
11. Companies using Hadoop for analytics bring meaning to satellite data // Skybox Imaging [Electronic resource]. – Mode of access : <http://www.skyboximaging.com/news/companies-using-hadoop-for-analytics-bring-meaning-to-satellite-data>. – Date of access : 08.05.2014.

12. Cascading // Concurrent, Inc. [Electronic resource]. – Mode of access : <http://www.cascading.org>. – Date of access : 08.05.2014.
13. HIPI: A Hadoop Image Processing Interface for Image-based MapReduce Tasks / C. Sweeney [et al.] [Electronic resource]. – Mode of access : http://cs.ucsb.edu/~cmsweeney/papers/undergrad_thesis.pdf. – Date of access : 08.05.2014.
14. HIPI: A Hadoop Image Processing Interface // Department of Computer Science at U. of VA [Electronic resource]. – Mode of access : <http://hipi.cs.virginia.edu>. – Date of access : 08.05.2014.
15. CUDA On Hadoop // The Apache Software Foundation [Electronic resource]. – Mode of access : <http://wiki.apache.org/hadoop/CUDA%20On%20Hadoop>. – Date of access : 08.05.2014.
16. jcuda.org [Electronic resource]. – Mode of access : <http://www.jcuda.org>. – Date of access : 08.05.2014.
17. He, C. CUDA performance study on Hadoop MapReduce Cluster / C. He, P. Du [Electronic resource]. – Mode of access : <http://www.slideshare.net/airbots/cuda-29330283>.
18. Mars: A MapReduce Framework on Graphics Processors / B. He [et al.] [Electronic resource]. – Mode of access : http://www.cse.ust.hk/gpuqp/Mars_tr.pdf. – Date of access : 08.05.2014.
19. Apache Spark // Spark Streaming Programming Guide [Electronic resource]. – Mode of access : <https://spark.apache.org/docs/0.9.0/streaming-programming-guide.html>. – Date of access : 08.05.2014.
20. What is Apache Mahout? // The Apache Software Foundation [Electronic resource]. – Mode of access : <https://mahout.apache.org>. – Date of access : 08.05.2014.
21. Apache Kafka // The Apache Software Foundation [Electronic resource]. – Mode of access : <http://kafka.apache.org/>. – Date of access : 08.05.2014.

A METHOD FOR COARSE SEGMENTATION OF PATHOLOGICAL LUNGS ON CT IMAGES USING RIBCAGE

V.A. Liauchuk, E.V. Snezhko

United Institute of Informatics Problems of the NAS of Belarus, Minsk

e-mail: vitali.liauchuk@gmail.com

A method is proposed for coarse segmentation of lung regions on CT images in the presence of lung pathology. The goal of the approach is to coarsely extract lung regions using only the skeleton (basically ribcage), thus eliminating most of the difficulties of pathological lung segmentation. The method is based on a model of elastic deformable surface, which tends to iteratively fit the ribcage. The method demonstrates good results in the task of segmenting images of lungs, strongly affected by a disease, what is important in terms of automated computer analysis of lung diseases.

Introduction

The problem of extraction of regions of interest is usually considered as a preprocessing step in various image analysis algorithms. In this context the task of lung segmentation on three-dimensional computed tomography (CT) images is of great importance in medical image analysis. A lot of work was done in order to achieve high-performance robust solutions for extracting lung regions on CT images [1–3]. In the case of healthy patients, lung segmentation on CT appears to be not so difficult due to large intensity differences between lung regions and surrounding tissue. Thus adaptive thresholding followed by morphological operations may lead to reasonable segmentation quality.

However in the case of pathology (e.g., pneumonia, fibrosis, tuberculosis, etc.) the task of extraction of correct lung regions becomes challenging. Being applied to images with pathology, conventional methods usually leave some lesion parts out of segmentation. Various approaches were developed in order to overcome the disadvantages of traditional lung segmentation methods [4–8]. In [4] the authors propose the segmentation method based on elastic registration of CT lung atlas to a target image of pathological lungs. In [7] and [8] the ribcage shape is used to set the initial iteration of the corresponding dynamic model (Snake Model and Active Shape respectively) which is used for further lung segmentation. Utilizing ribcage within the lung extraction process makes the corresponding algorithm more robust because ribcage is usually not affected by diseases and it can be easily extracted from images. In this regard, using ribcage for lung segmentation on CT images looks promising.

The main idea of the method proposed with this paper is to examine how well pathological lungs on CT images can be extracted using only the information about skeleton structure. Considering only high-density structures with intensity > 450 Hounsfield Units (HU) we can automatically avoid most of the challenges caused by the presence of pathology. In some cases lung disease leads to high-density structures inside the lung (e.g., calcifications). However, the proposed method appears to be insensitive to these minor changes of image structure. Thus, the goal of our method is to perform coarse segmentation of pathological lungs on CT images using only information on dense structures present on images (ribcage, spine and other bones). Focusing on analysis of lung pathology here we do not claim extraction of accurate lung regions. Instead we allow the final segmentation to capture some inner organs (trachea, heart, parts of stomach), expecting the method to segment most of the lesions.

In this paper, the segmentation is performed with use of energy-minimizing restricted active surface model similar to active contour. The shape is initialized near the ribcage and its evolution is guided by cost function considering internal and external energy. In contrast to conventional active contour, here external energy is defined by averaged bones intensity rather than gradient magnitude. The method is tested on CT images of pathological lungs taken from publically available image database [9].

1. Materials

The dataset used for testing our method consisted of 149 CT chest scans of tuberculosis patients. Three-dimensional CT images consisted of axial slices of size 512×512 pixels. The number of slices varied from 100 to 250 depending on scanning parameters and slice thickness (2.5 mm or 1.25 mm). In addition to original images (fig. 1, a) for each case three versions of segmentation were available: automated segmentation of lung regions based on adaptive thresholding and morphological operations (fig. 1, b), semi-automated lung segmentation (fig. 1, c) and manual segmentation of lesions and surrounding tissue (fig. 1, d). Semi-automated segmentation, which is considered as ground truth with this

study, consisted of two major steps: threshold-based automated segmentation followed by manual corrections of segmentation mistakes.

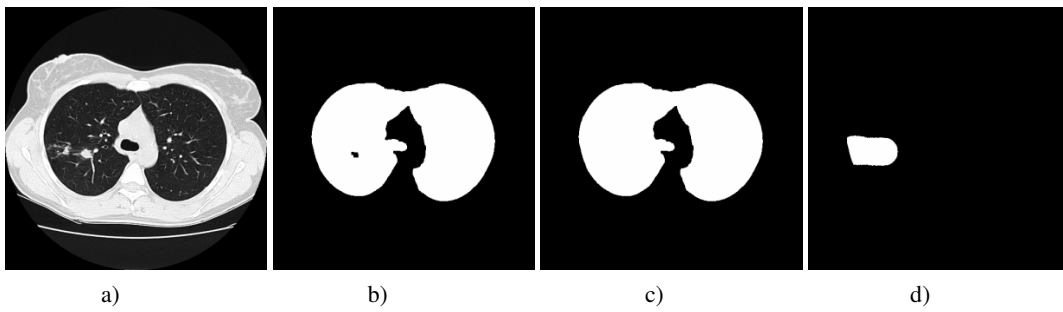


Fig. 1. Example slices of original image (a) and its segmented versions: automated threshold-based (b), semi-automated (c) and manual lesion segmentation (d)

2. Methods

In this work, a restricted deformable model is used for coarse segmentation of lungs using ribcage. The model incorporates a deformable surface represented by a set of vertices and edges. The evolution (deformation) of the surface is controlled by a cost function. By saying “restricted” we mean that each vertex is allowed to move towards only the predefined directions – along the line connecting the considered vertex and a reference point called “center”, which is not necessarily the geometric center of image (fig. 2, a). Thus, each vertex has only one degree of freedom instead of three. This restriction makes our model more stable and easy to control. However, it goes only under assumption that the shape of the region of interest to be extracted is simple enough.

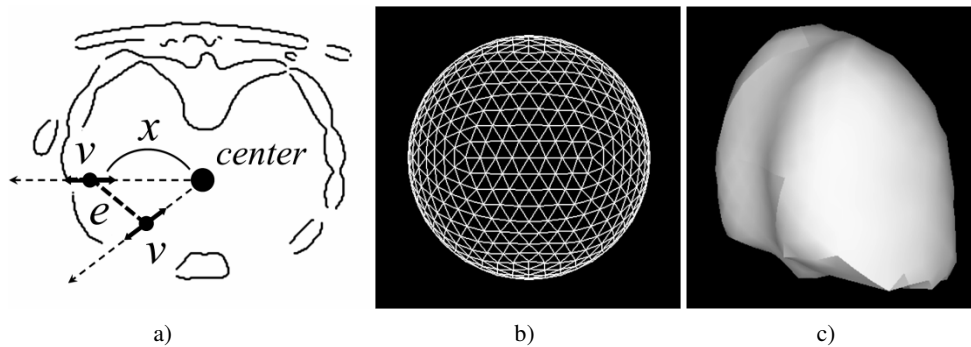


Fig. 2. Deformable surface: scheme of deformation (a), homogeneous spherical surface (b), deformed surface representing segmented lungs (d)

The proposed segmentation algorithm includes the following major steps:

- extracting bones, image preparation;
- generating homogeneous spherical surface;
- setting the initial iteration;
- iterative deformation of the surface;
- converting the final surface shape to raster mask.

Thus the final shape of the surface which best fits the ribcage while remaining smooth enough corresponds to the extracted lung region.

Image preparation step implies the extraction of high-density structures (ribcage, spine and other dense objects) present on the image and calculating the map of so called “closeness” to the ribcage for every image voxel. This is done by performing the following simple operations: binarizing the image using fixed threshold (450 HU), Gaussian smoothing (with STD = 6.5 mm), limiting the obtained values (to somehow equalize thick and thin bones’ contribution) and another Gaussian smoothing (smoothed field is suitable for gradient-descent minimization method). Applying the threshold of 450 HU almost completely eliminates the effects, appearing in the case of lung pathology (except for calcifications). With this preliminary study we didn’t try to remove any undesired high-density objects like bench or calcifications, what makes our current task a bit more difficult.

The shape of the active surface used for segmentation originates from *homogeneous sphere surface*, i.e. when not deformed all faces (triangles) of the surface should be equal. For this purpose we utilize the method of quasiregular tessellation of a sphere proposed in [10]. Thus, all the edges of the surface have almost the same length what makes our surface isotropic. In our case the obtained sphere surface contained 642 vertices connected by 1920 edges, which formed 1280 faces (fig. 2, b).

The cost-function to be minimized by the deformation process includes two terms which consider internal (elasticity) and external (“closeness” to ribcage) moving forces:

$$f(x) = f_{ex}(x) + \alpha \cdot f_{in}(x) = -\sum_{i=1}^n C(v_i(x)) + \alpha \cdot \sum_{j=1}^m \|e_j(x)\|^2, \quad (1)$$

where f_{in} and f_{ex} denote terms, corresponding to internal and external forces respectively, α – weighting coefficient, x is a vector of n arguments which control the positions of vertices, n and m being the numbers of vertices and edges respectively, v_i represents i -th vertice, e_j represents j -th edge, $C(v)$ denotes the value of closeness to ribcage and $\|e\|^2$ is the squared length of edge. Thus, our surface tends to get closer to the ribcage with minimum tension.

The *initial iteration* (initial surface) is set up in two steps. On the first step we coarsely find high-density structures by minimizing cost-function for each vertex separately with $\alpha = 0$ (eliminating the cost-function elasticity term). Then the coefficient α is set to the needed positive value and the procedure is repeated. The form of cost-function used with this study allows us to draw partial derivatives $\frac{\partial f}{\partial x_i}$, therefore finding x which minimizes the cost-function was performed via gradient descend method. Finally, image voxels which correspond to the interior of the deformed surface constitute the region of interest, i.e. coarsely segmented lung region. Example of the deformed surface is shown in fig. 2, c.

3. Experimental results

The lung segmentation method proposed was tested on the whole dataset of 149 CT scans. In different cases lungs were affected by disease at different extent, from small nodules to lung collapse and large fibrosis regions. The performance of our method was compared with an automated method based on thresholding and morphological operations (fig. 1, b) by means of calculating the following two indices. The first one is segmentation agreement which is helpful for general assessing of segmentation quality:

$$SA = \frac{2|\Omega \cap \Theta|}{|\Omega| + |\Theta|}, \quad (2)$$

where Ω corresponds to tested segmentation, Θ represents segmentation ground truth (semi-automated segmentation), \cap is intersection and $|\bullet|$ denotes regions size. The other index is important in terms of analysis of lung pathology and estimates the size of lesion part left outside of segmented region:

$$Miss = \frac{|\overline{\Omega} \cap \Delta|}{|\Theta|}, \quad (3)$$

where Δ corresponds to manually segmented lesions (fig. 2, d) and $\overline{\Omega}$ is exterior of Ω .

Fig. 3 demonstrates some segmentation results of the proposed method in comparison with threshold-based approach. Here we see that even significant changes in lung anatomy caused by diseases do not influence the segmentation results. However the method includes some mediastinum into the segmentation. Thus, segmentation agreement between the proposed method and ground truth segmentation is rather low (average 0.76 in contrast to 0.98 for threshold-based method).

Fig. 4 shows that traditional threshold-based method often leaves significant parts of affected lung regions outside of segmentation. With this respect the proposed method performs significantly better (better

in 87 cases out of 149), what is important for automated analysis of pathological lung images. Average lesion miss index for our method is 0.0038 in contrast to 0.061 for threshold-based one.

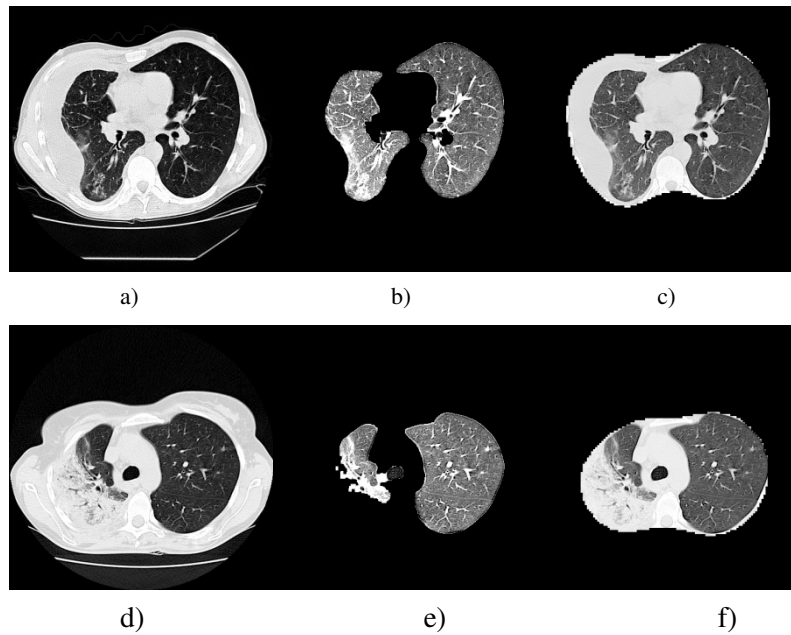


Fig. 3. Slices of original images (a, d), their threshold-based segmentation (b, e) and segmentation by the proposed method (c, f)

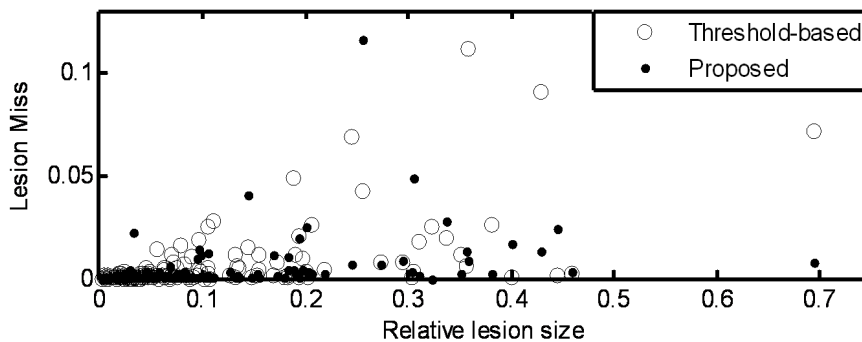


Fig. 4. The performance comparison of lung segmentation methods

Conclusion

The approach for pathological lung segmentation suggested with this study may be considered as a promising tool, useful for computer analysis of lung diseases. The segmentation result, produced by the method, is almost not influenced by any pathology.

However the method has disadvantages and its performance needs to be improved.

This work was partly funded by the National Institute of Allergy and Infectious Diseases, National Institutes of Health, USA through the CRDF project BOB1-31120-MK-13.

References

1. Armato, S.G. Automated lung segmentation for thoracic CT / S.G. Armato, W.F. Sensakovic // *Acad. Radiol.* – 2004. – Vol. 11 (9). – P. 1011–1021.
2. Automated lung segmentation in X-ray computed tomography / J.K. Leader [et al.] // *Acad. Radiol.* – 2003. – Vol. 10 (11). – P. 1224–1236.
3. Hu, S. Automatic lung segmentation for accurate quantitation of volumetric X-ray CT image / S. Hu, E.A. Hoffman, J.M. Reinhardt // *IEEE Trans. Med. Imag.* – 2001. – Vol. 20 (6). – P. 490–498.
4. Sluimer, I. Toward automated segmentation of the pathological lung in CT / I. Sluimer, M. Prokop, B. van Ginneken // *IEEE Trans. Med. Imag.* – 2005. – Vol. 24 (8). – P. 1025–1038.
5. Automatic segmentation of lung parenchyma in the presence of diseases based on curvature of ribs / M.N. Prasad [et al.] // *Acad. Radiol.* – 2008. – Vol. 15 (9). – P. 1173–1180.
6. Wang, J. Automated segmentation of lungs with severe interstitial lung disease in CT / J. Wang, Q. Li, F. Li // *Medical Physics.* – 2009. – Vol. 36 (10). – P. 4592–4599.

7. Meng, L. A New Lung Segmentation Algorithm for Pathological CT Images / L. Meng, H. Zhao // Proc. of Intern. Joint Conference on Computational Sciences and Optimization. – Sanya, 2009. – P. 847–850.
8. Robust Active Shape Model Based Lung Segmentation in CT Scans / S. Sun, C. Bauer, R. Beichel // Proc. of the 4th Intern. Workshop on Pulmonary Image Analysis. – 2011. – P. 213–223.
9. Belarus Tuberculosis Portal [Electronic resource]. – Mode of access: <http://tuberculosis.by> – Date of access : 21.01.2014.
10. Petrokovets, E. Quasiregular Tessellation of Arbitrary Sphere / E. Petrokovets, V. Kovalev // Proc. of the 5th Intern. Conference on Pattern Recognition and Information Processing (PRIP'99). – 1999. –Vol. 1. – P. 243–248.

STUDYING THE ADVANCED SEGMENTATION METHODS WITH THE COMPUTER-SIMULATED IMAGES

Y. Lisitsa¹, M. Yatskou¹, V. Apanasovich¹, T. Apanasovich²

¹Belarusian State University, Minsk;

²Kimmel Cancer Center at Thomas Jefferson University, USA

e-mail: lisitsa@bsu.com

The image simulation model is developed to simulate the three-channel fluorescence microscopy acquisition systems. The model is used to compare two advanced image analysis methods, realized in the software packages CellProfiler and CellDataMiner. The developed model allows to simplify and to make the best selection of the segmentation method to process specific fluorescent images of cancer cells in three-channel experiments. The CellDataMiner method shows more sufficient results than CellProfiler method for object localization and segmentation accuracy.

Introduction

Immunohistochemistry (IHC) is a well-established imaging technique that exploits intensity of staining dyes in tissue images to quantify the protein activity related to cancer development [1]. Even though there are a lot of flexible tools for automated analysis, the complexity and diversity in microscopic image data makes the developing suitable data analysis workflows as a challenging task [2].

The corn stone in the development in any image analysis technology is the need to study the modern and uprising algorithms thoroughly on a broad set of the reference experimental data, which are, in most cases, technically difficult or impossible to obtain. This problem can be successfully solved using synthetic images.

In this paper we present the simulation model developed to reproduce the fluorescent images from the three-channel acquisition microscopy systems with aim to compare efficiency of two advanced image analysis methodological approaches realized in the software packages CellProfiler and CellDataMiner [3, 4].

1. Protocol

The images, we consider, are three-channel slides (RGB colors), where the red channel is reserved for the protein *estrogen receptor* labeled by cyanine 5 dye (Cy5, $\lambda_{exc} = 654$ nm, $\lambda_{em} = 673$ nm). The red stain reveals the protein activity in cancer nuclei. The blue channel is reserved for the *emission* of molecule *DAPI, representing the DNA area* (DAPI, $\lambda_{exc} = 358$ nm, $\lambda_{em} = 461$ nm). Information on cytoplasm of *cancer cells* are collected in the green channel and formed by strong labeling Cy3-conjugated secondary antibody (Cy3, $\lambda_{exc} = 493$ nm, $\lambda_{em} = 518$ nm) [5].

2. Simulation model

The flowchart of the fluorescent image simulation integrates the simulation of the ideal image and its distortion with the effects, which appear during the image acquisition process [6]. The flowchart allows to simulate the cell population, where the each cell consists of two elements, where they are the cytoplasm and nucleus. As result the simulation of the single cell is the representation of its shape and intensity structure of cytoplasm and nucleus.

Assume that the shape of the cell elements is determined with the circle, where its radius is said with the normal distribution. Hence the pixel coordinates of the nucleus edge are described with the equation 1

$$x_{ij}^n = r_i^n \cdot \cos(\theta_j) \text{ and } y_{ij}^n = r_i^n \cdot \sin(\theta_j), \quad (1)$$

where j is the index of the pixel on an edge, $j=1,2,\dots,K$, K is the number of the pixels on an edge, θ_j – the angular coordinate for the j -th pixel, i is the index of the current cell, r_i^n is the radius of the i -th cell.

The parameter β is used to simulate the edge of the cytoplasm, it allows to vary its radius during the construction. The coordinates of the cytoplasm are determined with the following equation

$$x_{ij}^c = \xi \cdot r_i^c \cdot \cos(\theta_j) \text{ and } y_{ij}^c = \xi \cdot r_i^c \cdot \sin(\theta_j), \quad (2)$$

where j is the number of the pixel on the edge, $j=1,2,\dots,K$, K – is the number of pixels on the cytoplasm edge, θ_j is the value of the angular coordinate for the j -th pixel, i is the number of cell, ξ is the uniform distributed random variable at $[-\beta,\beta]$.

In order to simplify the simulation, the fluorescent intensity is assumed to be said with normal distribution $N(\mu = 1/2, D = 1/6)$ and it does not depend on the remoteness from the center of the object. As result the pixel intensities are distributed among 0 and 1.

The parameter n_c controls the number of clusters. The localization of the clusters is determined randomly and it is assessed with centers of the clusters. The cells, that belongs to the cluster, are normally distributed around it with the normal distribution $N(0,\sigma_c^2)$. The coordinates of the i -th cell are determined by

$$X_i = X_k + z \text{ and } Y_i = Y_k + z, \quad (3)$$

where X_k, Y_k are coordinates of the k -th cluster, z is the realization of the random variable with the normal distribution $N(0,\sigma_c^2)$.

As the cell populations can be compactly distributed in the image, they may be overlapping. The overlapping is estimated as

$$L_{lm} = \frac{|R_l \cap R_m|}{|R_l|}, \quad l \neq m, \quad (4)$$

where R_l, R_m are the areas of the pixels occupied with l -th and m -th cell accordingly, $l,m=1,2,\dots,K$. If L_{lm} is more than threshold value L^* , the new location for the m -th cell is determined [4].

The main distortions are caused by the photomultipliers at the acquisition system or by the three-dimensional structure of the specimen. The images of the objects that are not at the focal plane of the microscope are registered degraded. The Point Spread Function (PSF) is defined to simulate the degradation of the object registration and it is said with the equation 5 [6]:

$$PSF(x, y; a, b) = \exp\left(-\frac{0,5}{\sigma_g^2(a, b)} * [(x - a)^2 + (y - b)^2]\right), \quad (5)$$

where the level of the degradation is controlled by the value $\sigma_g^2(a, b)$, a and b are the center of the PSF. The noise from the photomultipliers is simulated as the addition to the image I_g where the pixel intensities are said with the normal distribution $N_g(0, \sigma_d^2)$ to the initial image I .

As a result the simulated image is described with the equation 6:

$$\hat{I} = \sum_{a,b \in G} PSF(x, y; a, b)I(a, b) + I_g, \quad (6)$$

where G is the set of the pixels around the degraded pixel.

3. Image analysis methods

We studied two modern methods for localization and segmentation accuracy on the simulated and real fluorescence images of breast cancer cells. The performance of the CellDataMiner technique was compared with the known semi-automated technique, realized in the software CellProfiler [3].

Algorithms

The CellProfiler (CP) exploits the standard methods to segment objects. However, it does not reflect the local image heterogeneity. In order to segment the nuclei we used the mixture of Gaussian adaptive thresholding at blue image component. The clumped cells were separated with the watershed method by the intensity profile of the blue channel of the image [3].

The second method (CDM, taken from the CellDataMiner) is fully-automated [4]. It extracts the prior information about studied objects from the image heterogeneity. Firstly, it extracts the information about the objects during the tumor segmentation. Simultaneously the method determines the partition for the adaptive segmentation. Secondly, the method determines the nuclei in the image based on the standard methods of segmentation such as the Otsu adaptive thresholding, watershed segmentation and Voronoi diagram. Finally the morphology based method is applied to improve object edges [7].

Criteria

First the algorithms were tested to their localization properties by four features:

- 1) part of the missed objects ε_m , that reflects the relative number of missed objects for each image;
- 2) part over segmented ε_o , that assesses the relative number of over segmented objects during the over segmentation;
- 3) part of the clumped objects ε_c ;
- 4) part of the false segmented objects ε_f , that reflects the number of objects classified as nuclei when there are indeed not.

The total localization error reflects the sum of all localization errors

$$\varepsilon = \varepsilon_m + \varepsilon_o + \varepsilon_c + \varepsilon_f \quad (7)$$

The second criteria that characterizes the algorithm quality is a segmentation accuracy δ [3]. It is sufficient to exploit only for the correctly localized objects. For each segmented object it can be calculated as

$$\delta = \left(1 - \frac{\sum xor(R, A)}{\sum or(R, A)}\right), \quad (8)$$

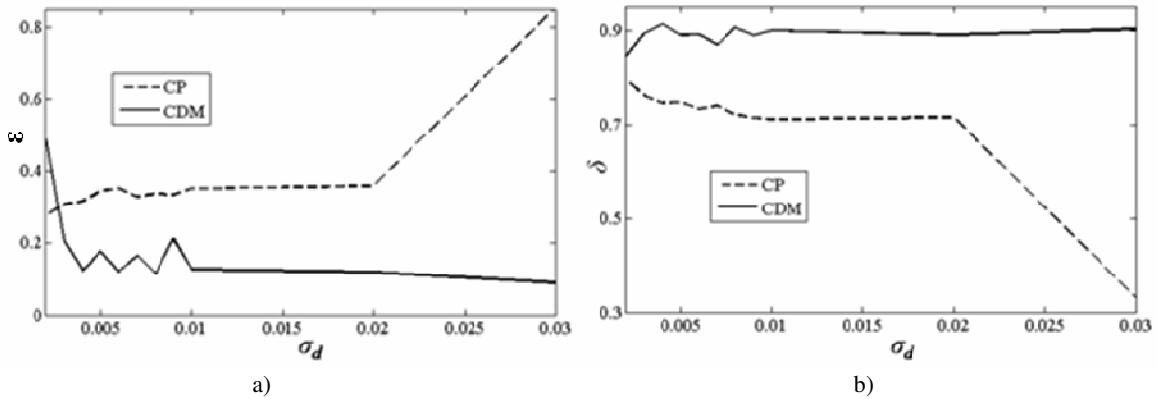
where R is a binary image, where pixels belonging to the reference manual nucleus are set to 1, and A is a binary image, where pixels enclosed by the nuclear boundary provided by the automated procedure are set to 1. As the each image contains about 600 validated objects it is sufficient to calculate the mean value M_δ of the estimates δ as a parameter for the whole image.

4. Results

The three sets of the images were simulated. The each set contains 38 images, where the parameter σ_d was varied among 0 and 0.3 (at arbitrary units).

The following parameters were set: $n_C = 2$, the number of cancer cells $n_{tum}=1200$, $L=0.1$, $5 \leq r_{tum}^n \leq 28$, the number of non-cancer cells $n_h = 1000$, $5 \leq r_h^n \leq 35$, the non-cancer cells do not overlap.

The results of the study for both methods are presented in figure and tables 1 and 2.



The total localization error ε (a) and mean segmentation accuracy $\langle \delta \rangle$ (b) for CP and CDM

According to the received data the CP allows to process data, where the σ_d is among 0 and 0.03, when the CDM works among 0.002 and 0.25. The total localization error ε for CP is from 0.3 to 0.6, on the other hand its value for CDM is from 0.1 to 0.2. The estimate ε_m forms the major contribute to the localization error for the CP (table 1). The segmentation accuracy decreases, when the noise increases, when the $\sigma_d=0.02$ the segmentation accuracy dramatically decreases (fig. b).

Table 1

The localization errors for CP					
CP					
σ_d	ε_m	ε_o	ε_c	ε_f	ε
0	0.359	0.010	0.050	0.004	0.424
0.002	0.198	0.011	0.062	0.007	0.278
0.005	0.252	0.007	0.080	0.007	0.346
0.007	0.227	0.011	0.085	0.006	0.329
0.009	0.238	0.007	0.083	0.008	0.336

0.010	0.257	0.013	0.076	0.006	0.352
0.020	0.254	0.004	0.095	0.007	0.360
0.030	0.540	0.002	0.000	0.312	0.855
0.040	∞	∞	∞	∞	∞

The ε_m , ε_o and ε_c increases as the noise value increases for CDM (table 2). The ε_f decreases, when σ_d is less than 0.004; if σ_d is less than 0.125 it reduces; and it dramatically increases when σ_d is more than 0.125. The mean segmentation accuracy $\langle \delta \rangle$ for CDM does not depend on the noise level along the working range and it is 20% higher than $\langle \delta \rangle$ for CP.

Table 2

The localization errors for CDM

CDM					
σ_d	ε_m	ε_o	ε_c	ε_f	ε
0	–	–	–	–	–
0.002	0.0032	0.0283	0.436	0.0049	0.4724
0.005	0.0052	0.0087	0.145	0.0018	0.1607
0.010	0.0043	0.0035	0.1146	0.0013	0.1237
0.050	0.0058	0.017	0.0569	0.0008	0.0805
0.100	0.0092	0.0084	0.0485	0.0007	0.0668
0.150	0.0061	0.0069	0.2455	0.0026	0.2611
0.200	0.0038	0.0084	0.08	0.001	0.0932
0.250	0.0254	0.0679	0.0133	0.0029	0.1095

Conclusion

The simulation model allows to objectively assess the localization and segmentation accuracy. It allows to study the peculiarities of the segmentation methods and to determine the working range for the studied methods. The fully-automated method realized at CellDataMiner shows more sufficient results for object localization and segmentation accuracy. Also it may be exploited at more broad noisy range than the method realized at CellProfiler.

References

1. Quantitative analysis of estrogen receptor heterogeneity in breast cancer / G.G. Chung [et al.] // Lab Invest. – 2007. – Vol. 87. – P. 662–669.
2. Sommer, C. Machine learning in cell biology – teaching computers to recognize phenotypes / C. Sommer, D.W. Gerlich // J Cell Sci. – 2013. – Vol. 9. – P. 4153–4163.
3. Improved structure, function and compatibility for CellProfiler: modular high-throughput image analysis software / L. Kamensky [et al.] // Bioinformatics. – 2011. – Vol. 8. – P. 1179–1180.
4. CellDataMiner is the software for data mining of fluorescent images of cancer cells / Y. Lisitsa [et al.] // Proc. of the 1th Intern. Conference “Means of Medical Electronics and New Medical Technologies” (Medelectronics’2012). – Minsk, 2012. – Vol. 1. – P. 67–70.
5. RB-pathway disruption in breast cancer: differential association with disease subtypes, disease-specific prognosis and therapeutic response / A. Ertel [et al.] // Cell Cycle. – 2010. – Vol. 9. – P. 4153–4163.
6. Computational framework for simulating fluorescence microscope images with cell populations / A. Lehmussola [et al.] // IEEE transactions on medical imaging. – 2007. – Vol. 26. – P. 1010–1016.
7. Fully-automated segmentation of tumor nuclei in cancer tissue images / Y. Lisitsa [et al.] // Proc. of the 1th Intern. Conference “Pattern Recognition and Information Processing” (PRIP’2011). – Minsk, 2011. – Vol. 1. – P. 116–120.

SIMULATION MODEL TO STUDY DENOISING METHODS

Y. Lisitsa, M. Yatskou, V. Skakun, A. Digris, I. Shingaryov, V. Apanasovich
Belarusian State University, Minsk
e-mail: ylisitsa@gmail.com

Fluorescence confocal microscopy combined with digital imaging constructs a basic platform for biomedical and pharmaceutic research. As the studies relying on analysis of digital images have become popular, the validation of image processing has become an important topic, where manual validation is impractical. Even though there are specific tools to simulate images from high-throughput microscopy systems based on fluorescent techniques, they are not applicable for the confocal image simulation. We present a simulation platform for generating synthetic confocal images of fluorescence-stained object with realistic properties.

Introduction

The Fluorescence Confocal Microscopy (FCM) allows the selection of a thin cross-section of the sample by rejecting the visual information coming from the out-of-focus planes. However, the small amount of energy radiated by the fluorophore and the huge light amplification performed by the photon detector to capture this visual information introduces a type of multiplicative noise described by a Poisson distribution [1].

High-complex image denoising tasks often make the implementation of static and predefined processing rules a cumbersome effort. The application of the respective algorithms to different markers or cell types then often requires parameter tuning or even re-programming of the software. Manual adaptations, however, are tedious and provide major obstacles, owing to the limited knowledge about the mathematics behind the image analysis algorithms [2].

With the simulated images, the validation can be carried more efficiently due to available ground-truth information. Simulation allows the validation with such a large amount of data not reasonable to achieve with the real experiments. There are a lot of flexible tools for the fluorescent image simulation; however they are not applicable for the FCM [3, 4].

1. Simulation pipeline

The image simulation is a multi-step process, where each step is involved in the image acquisition of the typical fluorescent microscopy experiment. First, the ideal image is generated. It contains the desired population of fluorescently labeled cells. Second it is degraded by the errors received from the measurement system.

The ideal grayscale image consisted in a big circular mask filled in a special texture together with or without two circular and square ideal objects.

The texture simulation is based on the Perlin noise [5]. The texture t in the location (x,y) is defined as

$$t(x, y) = B + \sum_{i=0}^{n-1} p^i \eta_{xy}(2^i), \quad (1)$$

where a weighted sum of n octaves of basic noise function $\eta_{xy}(\cdot)$, producing noise with given frequency, is calculated. The scaling of noise function is controlled with the persistence parameter p , and the bias with B . The different examples of the simulated structure are shown below.

The ideal image can contain the background simulated as the stripped pattern, where the length of strips is described with the normal distribution from L_{min} to L_{max} .

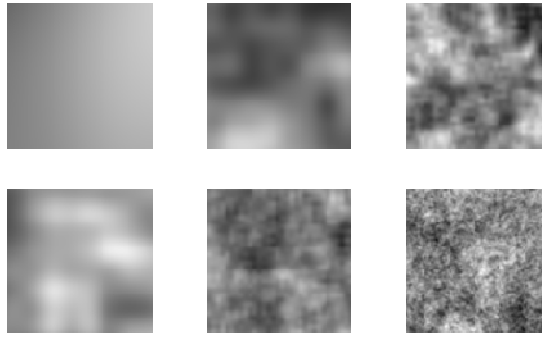


Fig. 1. Examples of the structure

Thereby the ideal image model is determined with the following parameters: the size of the image L , the radius of the circle R , the minimum intensity of the background I_{min}^b , the maximum intensity of the background I_{max}^b , the minimum length of the background strip L_{min} , the maximum length of the background strip L_{max} , persistence p , number of octaves n .

The ideal objects which can be added to the constructed image are the circle and the square, the diameter and side length are the half of R .

After the ideal image was simulated, it is corrupted with the Poisson noise. Before adding the noise the image is scaled in the range among 0 to $peak$. The noise is applied as

$$\begin{aligned}
 I(x, y) &= \alpha \cdot P(t(x, y)), \\
 \text{If } (I(x, y) < 0) &\Rightarrow I(x, y) = 0, \\
 \text{If } (I(x, y) > peak) &\Rightarrow I(x, y) = 255,
 \end{aligned} \tag{2}$$

where a discrete [random variable](#) P is said to have a Poisson distribution with parameter $t(x, y)$, α is a scaling factor.

2. Denoising methods

Nonlocal means (NLM) filter

The NL-means filter [6] is an evolution of the Bilateral filter which averages similar image pixels defined according to their local intensity similarity. The main difference between the NL-means and this filter is that the similarity between pixels has been made more robust to noise by using a region comparison, rather than pixel comparison and also that matching patterns are not restricted to be local. Thus, NLM filter is a nonlocal patch based filter. The idea of NL-means filter is based on the fact that every small window in a natural image has many similar windows in the same image.

Main algorithm has four parameters:

PatchHalfSize – size of the patch window f ,

WindowHalfSize – size of the neighborhood (research) window r ,

Sigma – standard deviation of Gaussian domain weights,

h –bandwidth (smoothing parameter) of range weights.

BM3D

The BM3D algorithm [7] uses grouping of neighborhoods whose surrounding square supersets have been found similar by a block-matching procedure. The data defined on these grouped neighborhoods is stacked together, resulting in 3D-data structures characterized by a high correlation along all the three dimensions. 3D-decorrelating transform, DCT and a 1D-orthonormal transform, and subsequent attenuation of the noise by spectrum shrinkage with hard-thresholding or Wiener filtering is then applied to 3D-structures. Inversion of the 3D-transform produces individual estimates for all grouped neighborhoods. These estimates are returned to their original locations and aggregated with other estimates coming from different groups.

The method has a number of adjustment parameters but almost all of them can be set to some predefined values. Therefore, it has one adjustable parameter: *Sigma* – standard deviation σ of noise.

Wiener filter

This filter filters the image using pixelwise adaptive Wiener filtering, using neighborhoods of size m -by- m to estimate the local image mean and standard deviation. The additive Gaussian noise is assumed in input data. The Wiener filter is the adaptive range filter and therefore should produce better results.

It has just one parameter:

WindowSize – size of the neighborhood window.

3. Denoising quality criteria

Let assume that $z_{ij}, y_{ij}, \hat{z}_{ij}$ are respectively pure, noisy and denoised pixels with coordinates i, j . For the Poisson statistics the following criteria can be used:

1. Maximum Likelihood Estimation for Poisson noise statistics (MLEP):

$$MLEP = \frac{2}{mn-1} \sum_{i=1}^m \sum_{j=1}^n x_{ij} \ln \left[\frac{x_{ij}}{z_{ij}} \right] - x_{ij} - z_{ij}, \quad (3)$$

where x_{ij} denotes either y_{ij} or \hat{z}_{ij} for noisy and denoised images respectively.

2. Structural Similarity Index (SSIM) (ranged from 0 to 1) [8].
3. MLE Poisson residuals between noisy and denoised images

$$residuals_{ij} = sign(x_{ij} - z_{ij}) * \sqrt{2(x_{ij} \ln(x_{ij} / z_{ij}) + z_{ij} - x_{ij})}, \quad (4)$$

where x_{ij} denotes either y_{ij} or \hat{z}_{ij} for noisy and denoised images respectively, function $sign(x)$ is equal to -1 in the case of negative x and 1 otherwise. In a case of perfect denoising residuals should not show any structure of the denoised image.

4. Visual (expert) comparison of denoised and pure initial image.

4. Simulated images

The simulated images can be exploited to compare algorithms for their effectiveness for image processing. Here there is the example of the usage of the simulated images to validate methods for their application to remove Poisson noise.

The example of the simulated image is shown in the fig. 2.

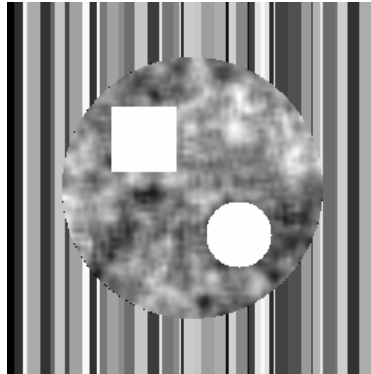


Fig. 2. The ideal image

The image is corrupted with the Poisson noise of peak is 22, α is 10. The intensity values are in the interval from 0 to 255. An example of the corrupted image is shown in the fig. 3.

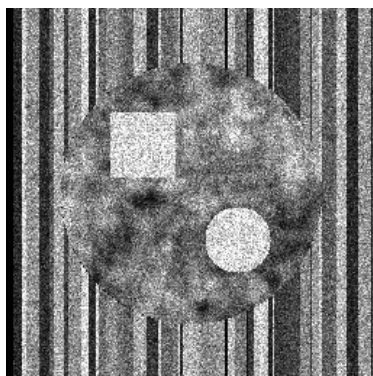


Fig. 3. Image corrupted with the Poisson noise

SSIM among pure image and noised image is 0.424, MLEP is 12.25.

Results

As the initial methods are developed for the Gaussian noise the anscombe transformation was used to adapt them for method application [9].

The result of denoising by NLMeans filter is shown on fig. 4, 5. Edges of ideal objects are blurred. The texture of the main object is not well restored and blurred.

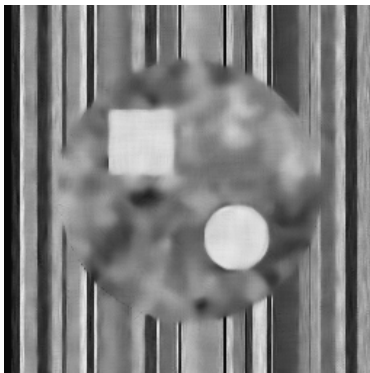


Fig. 4. Image restored with NLMeans

The image of residuals is full of strips from the background. The edges of the main objects are also visible.

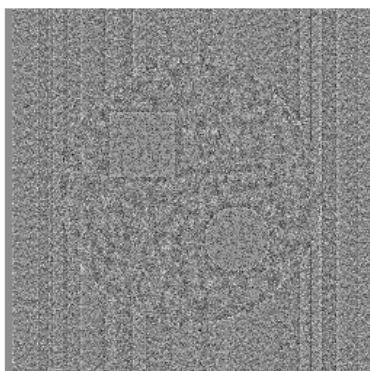


Fig. 5. The residuals for NLMeans method

The result of denoising by BM3D filter is shown on fig. 6, 7. All edges are clearly represented without artifacts and blurring. At the same time it should be noted that the texture within the big central object was not recovered well.

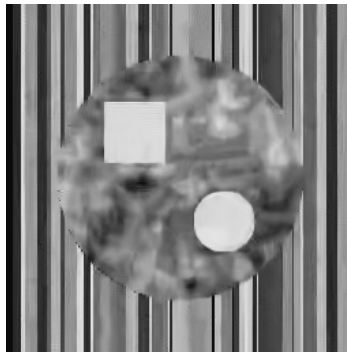


Fig. 6. Image restored with BM3D

The image of residuals contains well-visible strips from the dark parts of the original image.

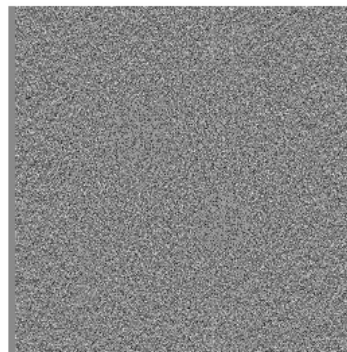


Fig. 7. The residuals for BM3D method

The result of denoising by wiener filter is shown on fig. 8, 9. The wiener filter passes some artifacts on the edges of the objects and the texture within the big central object was not recovered well.

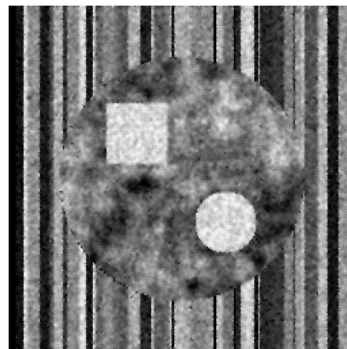


Fig. 8. Image restored with wiener filter

The image of residuals contains well-visible strips from the dark parts of the original image.

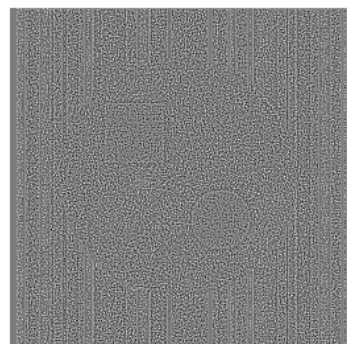


Fig. 9. The residuals for wiener method

The numerical results of the method validation are shown in the table. They correspond with the visual overview of denoising results.

Method\Metrics	MLEP	SSIM
NLM	3.5	0.786
BM3D	2.46	0.885
Wiener	5.54	0.695

The denoising methods reduce the MLEP in 3 times. The better results were received for the BM3D approach for both criteria. The worse results are received for the wiener filter. The numerical results are improved with the visual compare.

Conclusions

The usage of simulation data allows to determine the benefits and drawbacks of the exploited methods, that is not possible to extract by real data. It is suggested that the BM3D has more sufficient results at image denoising at it have less artifacts than NLMMeans and wiener filter.

References

1. Signal and noise modeling in confocal laser scanning fluorescence microscopy / G. Herberich [et al.] // *Med Image Comput. Assist Interv.* – 2012. – Vol. 15. – P. 381–388.
2. Sommer, C. Machine learning in cell biology - teaching computers to recognize phenotypes / C. Sommer, D.W. Gerlich // *J. Cell Sci.* – 2013. – Vol.
3. Computational framework for simulating fluorescence microscope images with cell populations / A. Lehmussola [et al.] // *IEEE Trans Med Imaging.* – 2007. – Vol. 26. – P. 1010–1016.
4. SimuCell: a flexible framework for creating synthetic microscopy images / S. Rajaram [et al.] // *Nat. Methods.* – 2012. – Vol. 9. – P. 634–635.
5. Perlin, K. An image synthesizer / K. Perlin // *SIGGRAPH Comput. Graph.* – 1985. – Vol. 19. – P. 287–296.
6. Buades, A. A Non-Local Algorithm for Image Denoising / A. Buades, B. Coll, J.-M. Morel // *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05).* – 2005. – Vol. 2. – P. 60–65.
7. Image Denoising by Sparse 3D-Transform-Domain Collaborative Filtering, *Image Processing* / K. Dabov [et al.] // *IEEE Transactions.* – 2007. – Vol. 16. – P. 2080–2095.
8. Image quality assessment: from error visibility to structural similarity / Z. Wang [et al.] // *IEEE Trans. Image Process.* – 2004. – Vol. 13. – P. 600–612.
9. Foi, A. Signal-dependent noise removal in Pointwise Shape-Adaptive DCT domain with locally adaptive variance / A. Foi, V. Katkovnik, K. Egiazarian // *Proc. 15th Eur. Signal Process. Conf., Eusipco.* – 2007.

BRAIN CODING OF INFORMATION ON THE OBJECT SHAPE BY ITS PHYSICAL MODEL

G. Losik¹, H. Brazevich²

¹United Institute of Informatics Problems of the NAS of Belarus, Minsk;

²Belarussian State Pedagogical University, Minsk

A physiological principle of pictorial information encoding in the brain is being considered in the given paper according to an informatics approach. It can be called as “place encoding”. Based on the theory on encoding and modeling the prototype, it has been proved that this principle that occurred in the brain is another unique method of encoding and decoding information. An encoded physical phenomenon is not replaced in the brain by a mathematical symbol, but by an analogous physical model. The given principle of encoding can be implemented in a person’s brain just due to the fact that while reproducing the genetic mechanism provides anthropomorphism. The physical similarity of a data medium (brain) structure characterized for all the representatives of the whole mankind.

Introduction

Material things which a person deals with, are differed in their size, surface texture, color and especially in their shape. People distinguish objects according to their shape very easily. However, the coding principles of three-dimensional objects shape, its memorization in the brain are still unknown [1, 2]. There is an unsettled question in cognitive psychology regarding a way how a visual image of a three-dimensional object, its shape is encoded in the brain. At the initial stages of visual information processing in sensory areas of the visual cortex the evaluation process concerning the length, slope, orientation, the lines thickness and curving, the angles, the outline closure takes place for a short time [3, 4].

1. Two appropriate tasks for analogous shape encoding

In accordance with our hypothesis, there are mechanisms of not digital but analogous processing of three-dimensional shape, their comparison, assessment of the different shapes similarity in the brain.

Studying these mechanisms based on the coding theory is considered to be a topical issue in connection with two important technical challenges: automatic images recognition presented in 3D dimension and the creation of “imaged” Internet. Automatic recognition of visual images has become more complicated while moving it from a two-dimensional representation of a recognizable object into three-dimensional representation of its shape.

2. The physical topology model of a 3D object shape

In natural sciences along with the widespread mathematical models a concept of a *physical* model is preserved. In the last case, some separate physical features of a prototype are carried from a prototype into a model, therefore the model can be considered as a “working” model. How a loss of information about a prototype can be understood while transferring from a physical to a mathematical model? It would seem that it is possible to transfer these data regarding a prototype and its environment into a digital model of an object. It is feasible but to some extent when it is not concerned to the degrees of freedom modeling, i.e. prototype kinematics, the degrees of freedom of its environment, flexibility, plasticity of their different parts, their aggregate states. Automatically a digital algorithmic mathematical model is to establish all combinations of degrees of freedom of a prototype initially.

However, the amplitude and direction of the degrees of freedom of an object and its parts can be varied so that a number of their combinations becomes endless. And in order to avoid it, the physical model intelligently saves only information about the alphabet and grammar of the compatibility of these degrees of freedom in the substance physics. It stores the information economically encoding it by “place” by means of the physical copy of topology, kinematics, a physical state of the prototype.

At the same time a disadvantage of physical model is that it is “matter-dependent”, in comparison with a digital, only algorithmic one. The physical model requires a physical media in the shape of strictly regulated substance for reproducing the recorded information in it. As opposed to it, the digital model is easily replicable; being an algorithm, it becomes “matter-independent”.

3. “Place-based” coding principle in the brain as a principle of physical modeling

While studying the processes of image information encoding in the brain, a psycho-physiological principle, so-called as “place-based coding” or “vector” encoding that was previously unknown, has been discovered [2, 4, 5]. In accordance with the approach of informatics, this principle is precisely can be considered as a coding principle by means of physical modeling. The existence of the principle has been proved experimentally in neuroscience by E. Sokolov’s scientific school. However, its essence remains unclear in terms of information theory. E. Sokolov’s principle of “place-based encoding in the brain” was introduced in contrast to the principle of “chain neurons encoding”, i.e. algorithm-based encoding. Processing of not all information perceived but only the image one is based on place encoding. A mechanism of place-based encoding is connected with the local neurons stimulation in the brain in specific local physical areas of the brain. Stimuli being perceived during a child’s development, depending on their different physical characteristics are taught by various detectors in neuronal layers to respond to the same stimulus, monotonous, in a strictly constant, but different parts of the brain.

Based on this principle, the location of the neuron in the space of sensory and motor cortex of the brain encodes the specifics of the information that it stores. Neurons-detectors that different depending on their place, form a “screen” of neurons or a *local analyzer*. Within its boundaries anatomical distance from one neuron to another stores the information about the degree of dissimilarity of the “quality” of the two types of information. While images training, a feature space is formed in the brain in a three-dimensional physical space of neurons according to which the image information that is stored in different neurons is varied.

Geometric interpretation of the local analyzer is a sphere, on the surface of which point detectors are located. The vectors from the center of the sphere produce azimuth angles between themselves towards detectors.

It was discovered experimentally that the distance-chords between the detectors on the surface of the sphere encode the degree of difference of *mental* characteristics of stimuli. And the azimuth angle between the vectors of detectors encodes the distance of physiological locations of these neuron-receptors in the brain. Thus, according to the principle of a place coding, an anatomical and morphological position of one neuron relative to the other one neuron in the brain creates a possibility of a vector encoding by means of an azimuth angle of the perceived degree of dissimilarity of different objects, including their shape.

An experimental method on how to determine the value of azimuth angle between the vectors of evoked potentials of the electroencephalogram for a particular person. A method on how to determine the value of the chords using the method of multidimensional scaling as well. Similarly chord value reflects a stimulus frequency tangling with one other. The beginning of the azimuth axes positions can be identified by the absence of a vector drift when a long-term adaptation of a local analyzer. Detectors of lower levels, characterized by slight features, raised, provide signals on the following screens of detectors, however, specialized in allocation of larger features, either the whole phenomena, such as a human face, geometric shapes, types of emotions, phonemes. Because of this there is a perception of objectivity.

The hypothesis can be considered as the assumption that the principle of place-based encoding is used while comparing objects with one other according to their *shape*. Such a principle of topological encoding of an object shape with a “neuron place” in the brain allows to estimate in a metrological way a degree of physiological arousal proximity of the “shape” of different neurons and, therefore, different geometric shapes that can be learned by the subject.

As a result, “the closeness of shape” because of anthropomorphic similarity of the brain in different individuals becomes measurable in a metrological way. This measurement is performed by means of the azimuth angle in a “spherical model” of the “neurons shape” field.

Conclusion

The paper shows that place-based coding is expressed in the creation of not a mathematical, but a *physical* model for the prototype. Place-based coding in the brain is a previously unknown unique case of information encoding and decoding in the brain. It is characterized by the fact the first signal spatio-temporal physical phenomena that are remembered by the brain, are encoded not metric scales, but a physical model. This principle is implemented in a person’s brain because its genetic mechanism, while reproducing, provides anthropomorphism, i.e. a physical resemblance to the structure of the brain as an information medium for all members of the human race.

The encoding principle by means of a “physical space” can be proved in the phylogenesis of the brain of the higher animals. It allows members of the same species, different individuals with the same anthropometry of the brain, according to the anthropometric criteria to compare a shape of various objects

in everyday use in an analogous and approximately in the same way to estimate dissimilarity of the topology of their shape.

References

1. Sokolov, E.N. Long-Term Memory, Neurogenesis, and Signal Novelty / E.N. Sokolov, N.I. Nezlina // *Neuroscience and Behavioral Physiology*. – 2004. – Vol. 34, no. 8. – P. 847–857.
2. Mangina, C.A. Neuronal plasticity in memory and learning abilities / C.A. Mangina, E.N. Sokolov // *International Journal of Psychophysiology*. – 2005. – November. – P. 14–29.
3. Sokolov, E.N. Neuronal models in the orienting reflex / E.N. Sokolov // *The Central Nervous System and Behavior*. Macy Foundation / ed. Brazier, M.A. – N.Y., 1960. – P. 187–275.
4. Sokolov, E.N. Perception and the conditioning reflex: vector encoding / E.N. Sokolov // *Int. J. Psychophysiol.* – 2000. – № 35. – P. 197–217.
5. Model of cognitive processes / E.N. Sokolov [et al.] // *Advances in psychological science. Biological and cognitive aspects*. – Hove : Psychological press, 1998. – Vol. 2. – P. 355–379.

FREQUENCY ANALYSIS OF GRADIENT DESCENT METHOD FOR IMAGE RESTORATION

A. Makovetskii¹, V. Kober²

¹Department of Mathematics, Chelyabinsk State University, Russia;

²Department of Computer Science, CICESE, Ensenada, B.C. 22860, Mexico
e-mail: artemmac@csu.ru

A function of two variables distorted by a known linear operator and additive noise is usually used in the problem of image restoration. It was shown that to solve the problem metrical as well as topological characteristics of a function of two variables should be used. In the paper with the help of frequency analysis we derive conditions when the gradient decent method is able to restore images. Computer simulation results are provided to illustrate the performance of the gradient decent method for restoration of uniformly blurred signals.

Introduction

Numerous papers on image restoration consider the restoration of functions distorted by a linear operator using different variations of the gradient descent method [1–4]. In this paper, analysis of the gradient descent method for image restoration is performed in order to find shortcomings of the method and to improve its performance using linear variations of a function.

Let \mathbf{u} be an original image, \mathbf{A} be a distorting compact operator, and \mathbf{u}_0 be a degraded image, that is, $\mathbf{u}_0 = \mathbf{A} \mathbf{u}$. Consider the restoration problem of the original image \mathbf{u} using known \mathbf{A} and \mathbf{u}_0 . Denote by $J(\mathbf{u})$ the following functional:

$$J(\mathbf{u}) = \| \mathbf{A} \mathbf{u} - \mathbf{u}_0 \|_{L_2}^2. \quad (1)$$

To restore the function \mathbf{u} , the following variation problem is stated:

$$\begin{aligned} \mathbf{u} = \operatorname{argmin} \| \mathbf{A} \mathbf{u} - \mathbf{u}_0 \|_{L_2}^2, \\ \mathbf{u} \in BV(\Omega), \end{aligned} \quad (2)$$

where Ω is the domain of the function \mathbf{u} in \mathbb{R}^2 . A common way of solving the problem in (2) is to use the gradient descent method [5]. The iterations of the gradient descent method are given as

$$\mathbf{u}_{k+1} = \mathbf{u}_k - \alpha_{k+1} \cdot J'(\mathbf{u}_k). \quad (3)$$

1. Linear variation for continuous and discrete function

The norm is a metrical characteristic of a function of two variables. Continuous functions of two variables also have a set of topological characteristics referred to as linear variations. Kronrod [6] introduced the notion of a regular component of the level set of a continuous function of two variables. The simplest topological characteristic in the linear variation theory is a number of regular components for all level sets of a function. Full information about linear variations of a function is contained in the one-dimensional tree of a function of two variables function is contained in the one-dimensional tree of a function of two variables.

Let $\Phi_{\mathbf{u}}(\tau)$ be the number of regular components of a level set \mathbf{t} for a continuous function [6]. The first Kronrod's linear variation is defined as

$$V(\mathbf{u}) = \int_{-\infty}^{+\infty} \Phi_{\mathbf{u}}(\tau) d\tau. \quad (4)$$

Let \mathbf{w} be a binary discrete function $\mathbf{w} = (w_{i,j})$, where $w_{i,j} \in \{0,1\}$, for all pairs (i,j) . A subset of such pairs (i,j) when $w_{i,j} = 1$ and all elements of the subset are connected by the 8-connectivity, is called the connected component of the binary function \mathbf{w} . For a number $k \in \mathbb{N}$ and a discrete function \mathbf{u} we define the following indicator function χ_k :

$$\chi_k(u_{i,j}) = \begin{cases} 1, & u_{i,j} \geq k \\ 0, & u_{i,j} < k \end{cases} \quad (5)$$

Definition 1. The number $V_k(u_{i,j})$ of connected components for a level k , $k \in \mathbb{N}$ of the discrete function \mathbf{u} is called the number of connected components of the binary discrete function $\chi_k(u_{i,j})$.

Definition 2. The linear variation $V(u_{i,j})$ of a discrete function \mathbf{u} is defined as follows:

$$V(u_{i,j}) = \sum_{k=0}^{+\infty} V_k(u_{i,j}). \quad (6)$$

Let us compute the discrete gradient $\nabla u_{i,j}$ of \mathbf{u} at (i,j) as

$$\nabla u_{i,j} = (u_{i+1,j} - u_{i,j}, u_{i,j+1} - u_{i,j}). \quad (7)$$

Suppose that if the pair (i,j) is outside of the domain of the function \mathbf{u} , then $u_{i,j} = 0$.

2. Analysis of gradient decent method in Fourier domain

Let $v \in L_2(\Omega)$ and $u_0 = \mathbf{A}v$ be original and degraded functions, respectively. Here the operator \mathbf{A} is a centered uniform blurring

$$\mathbf{A}v(t) = \frac{1}{2\Delta} \int_{t-\Delta}^{t+\Delta} v(\tau) d\tau, \quad (8)$$

where t is an arbitrary point of Ω , Δ is the parameter of blurring.

Remark. The operator \mathbf{A} is a self-adjoint operator, that is, $\mathbf{A}^* = \mathbf{A}$.

The function $\text{rect}(t)$ is defined as follows:

$$\text{rect}(t) = \begin{cases} 1, & -\frac{1}{2} \leq t \leq \frac{1}{2} \\ 0, & \text{otherwise} \end{cases} \quad (9)$$

Suppose that the function $v(t)$ is defined in the interval

$$v(t) = \begin{cases} h_w - \frac{T_b}{2} \leq t \leq \frac{T_b}{2} \\ 0, & \text{otherwise} \end{cases} \quad (10)$$

The functions $v(t)$ and $\text{rect}(t)$ are related by the expression

$$v(t) = h_v \cdot \text{rect}\left(\frac{t}{T_v}\right). \quad (11)$$

We also consider the following function $a(t)$:

$$a(t) = \begin{cases} \frac{1}{2\Delta}, & -\Delta \leq t \leq \Delta \\ 0, & \text{otherwise} \end{cases}. \quad (12)$$

The functions $a(t)$ and $\text{rect}(t)$ are related by:

$$a(t) = \frac{1}{2\Delta} \cdot \text{rect}\left(\frac{t}{2\Delta}\right). \quad (13)$$

By applying the operator \mathbf{A} to the function v , we obtain the convolution between the functions $v(t)$ and $a(t)$,

$$\begin{aligned} u_0(t) = v(t) * a(t) &= \int_{-\infty}^{+\infty} v(\tau) \cdot a(t - \tau) d\tau = \frac{1}{2\Delta} \int_{-\infty}^{+\infty} v(\tau) \cdot \text{rect}\left(\frac{t}{2\Delta} - \tau\right) d\tau = \\ &= \frac{1}{2\Delta} \int_{-\infty}^{+\infty} v(\tau) \cdot \text{rect}\left(\frac{\tau}{2\Delta} - \tau\right) d\tau = \frac{1}{2\Delta} \int_{t-\Delta}^{t+\Delta} v(\tau) \cdot 1 d\tau. \end{aligned} \quad (14)$$

Therefore

$$u_0(t) = v(t) * a(t) = h_v \frac{1}{2\Delta} \left(\text{rect}\left(\frac{t}{T_v}\right) * \text{rect}\left(\frac{t}{2\Delta}\right) \right) \quad (15)$$

Consider the Fourier transform $U_0(f) = \mathfrak{F}(u_0(t))$ of the function $u_0(t)$. On the basis of the following properties of the Fourier transform:

$$\mathfrak{F}(\text{rect}(t)) = \text{sinc}(f), \quad (16)$$

$$\mathfrak{F}\left(u\left(\frac{t}{T}\right)\right) = TU(Tf), \quad (17)$$

and the convolution theorem, we get

$$\begin{aligned} U_0(f) = \mathfrak{F}(u_0(t)) &= h_v \frac{1}{2\Delta} T_v 2\Delta \left(\text{sinc}(T_v f) \cdot \text{sinc}(2\Delta f) \right) = \\ &= h_v \frac{1}{2\Delta} T_v 2\Delta \left(\frac{\sin(T_v f)}{T_v f} \cdot \frac{\sin(2\Delta f)}{2\Delta f} \right) = h_v \frac{1}{2\Delta} \left(\frac{\sin(T_v f) \sin(2\Delta f)}{f^2} \right). \end{aligned} \quad (18)$$

It is obvious that the uniform blurring acts as a low-pass filter, that is, high frequency components of the functions $U_0(f)$ are suppressed.

For the operator \mathbf{A} defined in eq. (8), the functional $J'(u_k)$ takes the following form:

$$J'(u_k) = 2 A^*(A u_k - u_0) = 2 A^2 u_k - 2 A u_0. \quad (19)$$

Eq. (3) becomes

$$u_{k+1} = u_k - \alpha_{k+1} \cdot J'(u_k) = u_k - 2\alpha_{k+1} A^2 u_k + 2\alpha_{k+1} A u_0. \quad (20)$$

Suppose that the parameters of the gradient descent method are the same for all k , that is,

$$\alpha_n = \alpha, n = 1, \dots, k + 1. \quad (21)$$

Proposition 1. The function u_{k+1} , obtained at the iteration $k + 1$ of the gradient descent method can be expressed as follows:

$$u_{k+1} = ((1 - 2\alpha A^2)^{k+1} + ((1 - 2\alpha A^2)^k + \dots + (1 - 2\alpha A^2)^1 + 1) 2\alpha A) (u_0). \quad (22)$$

Proposition 2.

$$\mathfrak{F}(u_{k+1}) = (1 - (1 - \text{sinc}(2\Delta f)) (1 - 2\alpha \text{sinc}^2(2\Delta f))^{k+1}). \quad (23)$$

Therefore, the original function cannot be restored by the gradient descent method with any number of iterations if and only if these zeros do not coincide with zeros of the original function. In contrary, if the spectrum of the original function contains zeros and these are located in the same frequency positions that those of the degradation function spectrum then the function u_k converges pointwise to the function v . Note, that if the gradient decent method is not able to completely restore the original function than the projection method [7] based on linear variations could be utilized. The projection method helps to restore pretty well the values of local function extrema at know positions.

3. Computer Simulation

Let us define the following discrete function shown in (fig. 1):

$$v_i = \begin{cases} 0, & 0 \leq i \leq 94 \\ 210, & 95 \leq i \leq 105 \\ 0, & 106 \leq i \leq 200 \end{cases}. \quad (24)$$

Denote by $d(v, u_k)$ the distance in the considered functional space between functions v and u_k in the metric L_2 . The computed distances are as follows: $d(v, u_{200,000}) = 63.31$ and $d(v, u_{250,000}) = 63.31$. Hence, the restoration error is not reduced versus the number of iterations of the gradient decent method.

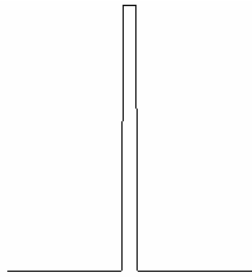


Fig. 1. Test discrete signal

The restored function and difference between the original and restored functions are shown in (fig. 2) and (fig. 3), respectively.

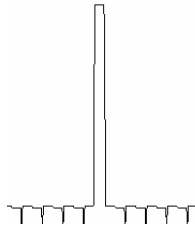


Fig. 2. Restored discrete signal for the uniform blur of $\Delta=10$



Fig. 3. Difference between the original and restored signals for the uniform blur of $\Delta=10$

Conclusion

In this paper the help of spectral analysis the performance of the gradient decent method for restoration of images blurred by a linear operator was analyzed. We derived conditions when the gradient decent method is able to restore images. Computer simulation results are provided to illustrate the performance of the gradient decent method for restoration of uniformly blurred signals. The performance of the gradient decent method can be improved by using linear variations and early proposed projection method.

The work is partially supported by RFBR grant № 13-01-00735.

References

1. Rudin, L.I. Total variation based image restoration with free local constraints / L.I. Rudin, S. Osher // Proc. Intern. Conf. Image Processing. – 1994. – Vol. 1. – P. 31–35.
2. Chambolle, A. Image recovery via total variational minimization and related problems / A. Chambolle, P.L. Lions // Numer. Math. – 1997. – Vol. 76. – P. 167–188.
3. An iterative regularization method for total variation based image restoration / S. Osher [et al.] // MultiscaleModelling and Simulation. –2005. – Vol. 4. – P. 460–489.
4. Chambolle, A. An Algorithm for Total Variation Minimization and Applications / A. Chambolle // Journal of Mathematical Imaging and Vision. – 2004. – Vol. 20. – P. 89–97.
5. Snyman, J.A. Practical Mathematical Optimization: An Introduction to Basic Optimization Theory and Classical and New Gradient-Based Algorithms / J.A. Snyman. – Springer Publishing, 2005.
6. Kronrod, A. On functions of two variables / A. Kronrod // Uspehi Mat. Nauk. – 1950. – Vol. 1, no. 35. –P. 24–134.
7. Makovetskii, A. Image restoration based on topological properties of functions of two variables / A. Makovetskii, V. Kober // Proc. SPIE Applications of Digital Image Processing XXXV. – 2012. – Vol. 8499. – P. 84990A.

HeNLM-3D: A 3D COMPUTER TOMOGRAPHY IMAGE DENOISING ALGORITHM³

N. Mamaev, A. Lukin, D. Yurin
Lomonosov Moscow State University, Russia
e-mails: nikolay.mamaev93@mail.ru, lukin@ixbt.com, yurin@cs.msu.su

A 3D filtering algorithm for CT images is proposed (HeNLM-3D). It is an extension of the known Local Jets method (LJNLM-LR) and our previous 2D algorithm (HeNLM). The proposed algorithm is based on the expansion of pixel neighborhoods into Hermite functions which form the orthonormal system. Tests on real CT images have shown that HeNLM-3D performs filtering better in high-detailed areas in comparison with LJNLM-LR and classical NLM method.

Introduction

Computed tomography (CT) is a method for studying the internal structure of the body by using a Radon transform [1]. Noise is always present in obtained CT images, and it becomes higher when the X-ray radiation dose is reduced. So, the effective noise reduction in CT images allows reducing the radiation dose [2, 3]. Noise in CT images is close to Gaussian [4]. CT images comprise the series of slices which forms a 3D image. It is important to note that these slices are a feature of the registration procedure, not of the object itself. For example, a blood vessel is not required to lie in a plane of the slice or to be perpendicular to it. Obviously, in this case the preferred direction of filtering should be the local direction of the vessel, not a slice plane or a perpendicular to the slice. Therefore, fully 3D algorithms – isotropic or with locally adaptive anisotropy – are needed.

Today some of the most effective image denoising algorithms are those in which each pixel of the output image is formed as a weighted sum of pixels of the source image. At the same time the weights depend on the similarity of whole pixels neighborhoods (patches) [5–7]. In a non-local means (NLM) algorithm [5] weights depend on the Euclidian distance between patches of pixels. LJNLM-LR [6] and GFNLM [7] algorithms are the extension of [5] and weights depend on the Euclidian distance between the feature vectors which characterize the patches. In [6] components of the feature vector are values of Taylor series expansion coefficients, which, in turn, are the values of image convolution with derivatives of the Gaussian function. One of advantages of [6] is the invariance of features to rotation. In [7] features are based on Gabor functions.

In this article a 3D method for noise reduction is proposed based on Hermite functions expansion which forms the orthonormal system. Inheriting all advantages of [6] the HeNLM-3D method is fully three-dimensional. It can better distinguish textures due to higher independence of feature vector components and better description of high-frequency components of the pixel neighborhood.

1. Hermite Functions

Hermite functions of order n are defined as

$$\psi_n(x) = \frac{1}{c_n} e^{-\frac{x^2}{2}} \frac{d^n(e^{-x^2})}{dx^n} = \frac{(-1)^n}{c_n} H_n(x) e^{-\frac{x^2}{2}}, \quad (1)$$

where $c_n = \sqrt{\sqrt{\pi} 2^n n!}$, $H_n(x) = (-1)^n \frac{d^n e^{-x^2}}{dx^n} e^{x^2}$ is a Hermite polynomial.

They form the orthonormal system in $L_2(-\infty, +\infty)$ [8]:

$$\int_{-\infty}^{+\infty} \psi_n(x) \psi_m(x) dx = \delta_{nm}. \quad (2)$$

Multiscale Hermite functions are defined as follows:

³ The research was supported by the RFBR grant 13-07-00584.

$$\psi_n^\sigma(x) = \frac{1}{\sigma} \psi_n\left(\frac{x}{\sigma}\right). \quad (3)$$

The multiplier $1/\sigma$ in (3) is introduced for similarity of the filter response to similar patterns across different scales.

Some of the first Hermite functions and derivatives of the Gaussian function are shown in the Fig. 1. We can see that unlike Gaussians, Hermite functions have approximately the same overall amplitude and higher values at the sides of their support interval.

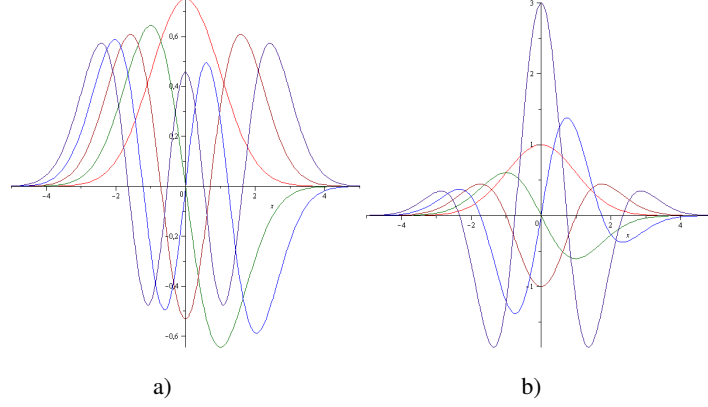


Fig. 1. Hermite functions (a) and derivatives of the Gaussian function (b)

We define 3D Hermite functions as

$$\psi_{nmk}^\sigma(x, y, z) = \psi_n^\sigma(x) \psi_m^\sigma(y) \psi_k^\sigma(z). \quad (4)$$

2. 3D Hermite Functions based Non-Local Means

In [9] we have proposed a modification of LJNLM-LR algorithm [6] based on Hermite functions instead of derivatives of the Gaussian function. In this article, we propose a 3D extension of [9] (HeNLM-3D), i.e. our pixel neighborhood is now a 3D cube and elements of the feature vector are values of a convolution of the source image series with 3D Hermite functions:

$$h_{nmk}^\sigma(x, y, z) = I(x, y, z) * \left(\frac{1}{N} \psi_{nmk}^\sigma(x, y, z) \right), \quad (5)$$

where N is the number of different Hermite functions of $n+m+k$ order.

As well as in [6], the obtained features are transformed to the new coordinate system (see (6)) (ξ_1, ξ_2, ξ_3) for invariance to rotation. In [6] and [9] the new 2D coordinate system was chosen as a gradient direction and the normal to it. In the case of a 3D coordinate system, two directions have to be chosen. Therefore, the transfer of rotation formulas from methods [6], [9] is impossible. In this article we propose using the eigenvectors of a structure tensor [10] as a new coordinate system:

$$M = \left\langle g^\sigma \cdot g^{\sigma T} \right\rangle, \text{ where } g^\sigma = (g_x^\sigma, g_y^\sigma, g_z^\sigma)^T \text{ is the image gradient.} \quad (6)$$

Averaging in (6) means averaging of matrix elements by a Gaussian filter with $\sigma_M = 3\sigma$ which is 3 times higher than σ used for the derivatives computation. It is important to note that, for example, in a small neighborhood of the vessel with a radius on the order of σ , the eigenvector corresponding to the minimal eigenvalue is directed along the vessel axis which has a favorable effect on image filtering.

To obtain an explicit expression for components of the feature vector in the new coordinate system, we consider two coordinate systems: the original one (x_1, x_2, x_3) related to rows, columns and slices of the 3D image, and the rotated coordinate system (ξ_1, ξ_2, ξ_3) :

$$\begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix} = R \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad R = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, \quad a_{ij} = (\rho_i, \xi_j), \quad (7)$$

where $i, j = 1, 2, 3$, ρ_i, ξ_j are basis unit vectors of basis $(x_1, x_2, x_3) \equiv (x, y, z)$ and (ξ_1, ξ_2, ξ_3) respectively.

Then the differential operator in this new coordinate system can be expressed through the linear combination of operators in the old coordinate system:

$$\frac{d}{d\xi_i} = \sum_{j=1,2,3} R_{ij} \frac{d}{dx_j}. \quad (8)$$

Therefore the expression for the Gaussian function derivatives can be written as

$$\begin{aligned} \frac{dG_\sigma}{d\xi_1^n d\xi_2^m d\xi_3^k} &= \left(\sum_{j=1}^3 R_{1j} \frac{d}{dx_j} \right)^n \left(\sum_{j=1}^3 R_{2j} \frac{d}{dx_j} \right)^m \left(\sum_{j=1}^3 R_{3j} \frac{d}{dx_j} \right)^k G_\sigma = \\ &= \left(\sum_{i=0}^n \sum_{j=0}^i C_{i-j,j}^n \left(R_{11} \frac{d}{dx_1} \right)^{n-i} \left(R_{12} \frac{d}{dx_2} \right)^{i-j} \left(R_{13} \frac{d}{dx_3} \right)^j \right) \cdot \\ &\quad \left(\sum_{i=0}^m \sum_{j=0}^i C_{i-j,j}^m \left(R_{21} \frac{d}{dx_1} \right)^{m-i} \left(R_{22} \frac{d}{dx_2} \right)^{i-j} \left(R_{23} \frac{d}{dx_3} \right)^j \right) \cdot \\ &\quad \left(\sum_{i=0}^k \sum_{j=0}^i C_{i-j,j}^k \left(R_{31} \frac{d}{dx_1} \right)^{k-i} \left(R_{32} \frac{d}{dx_2} \right)^{i-j} \left(R_{33} \frac{d}{dx_3} \right)^j \right) G_\sigma, \end{aligned} \quad (9)$$

where $C_{pq}^n = \frac{n!}{p!q!(n-p-q)!}$.

A 3D Hermite function is a product of the Gaussian function derivative, the multiplier $\exp(-(x^2 + y^2 + z^2)/2)$ and the coefficient c_i . Therefore, from (7) and the radial symmetry of the exponential function it follows that

$$c_n c_m c_k \tilde{\psi}_{nmk}^\sigma = \sum_{i=0}^r \sum_{j=0}^i \alpha_{r-i,i-j,j} c_{r-i} c_{i-j} c_j \psi_{r-i,i-j,j}^\sigma, \quad \text{where } r = n + m + k$$

$$\alpha_{abc} = \sum_{i=\max\{0,a-n\}}^{\min\{a,m+k\}} \sum_{j=\max\{0,a-n-m\}}^{\min\{i,k\}} \sum_{p=\max\{0,b-n, a-n-i-b\}}^{\min\{b,m+k\}} \sum_{q=\max\{0,b-n-m, p-m+i-j\}}^{\min\{p,k-j\}} R_{11}^{a-i} R_{21}^{i-j} R_{31}^j R_{12}^{b-p} \cdot \quad (10)$$

In (10) $\tilde{\psi}_{nmk}^\sigma$ is the Hermite function in a new coordinate system. We note that formula (10) suits for

Gaussian function derivatives by setting coefficients $c_i = 1$.

Finally, the feature vector is determined as follows:

$$\mathbb{V} = \{ \tilde{h}_{nmk}^\sigma ; n + m + k \leq r, \sigma \in S \}. \quad (11)$$

Here r is the maximal order of the Hermite function, and S is a set of different scales. And the value of the output pixel (voxel) $f(\mathbf{p})$ is the weighed sum of source image pixels (voxels) $I(\mathbf{p})$ from the neighborhood Q :

$$f(\mathbf{p}) = \frac{1}{\sum_{\xi \in Q} w(\mathbf{p}, \xi)} \sum_{\xi \in Q} w(\mathbf{p}, \xi) I(\mathbf{p} + \xi). \quad (12)$$

The weights depend on the similarity of feature vectors (11):

$$w(\mathbf{p}, \xi) = \exp\left(-\frac{\|\mathbf{v}(\mathbf{p}, \xi) - \mathbf{v}(\mathbf{p} + \xi)\|_2^2}{2\rho^2}\right). \quad (13)$$

The use of Hermite functions instead of derivatives of the Gaussian function provides better characterization the pixel neighborhood because their orthogonality means less dependence between feature vector components and, accordingly, their greater significance. In addition, the localization region is expanded (see fig. 1) and the peripheral data of a local neighborhood is better taken into account.

3. Results

Fig. 2 shows the source CT image, and fig. 3 shows the enlarged fragment of it and the results of filtering by two methods: LjNLM-LR and HeNLM-3D. The size of a neighborhood for LjNLM-LR method has been set to 21×21 pixels. Because in our CT images the physical size of a voxel along x and y axes is about 2.1 times less than its size along z axis, we set the size of a neighborhood for the HeNLM-3D method to $21 \times 21 \times 9$ voxels. For the same reason, the parameter σ for Gaussian and Hermite functions has been set to 2.1 for x and y axes and 1.0 for z axis. The maximal order of Hermite functions and Gaussian function derivatives has been set to 4. The parameter ρ for all methods has been manually adjusted, so that PSNR between noisy and filtered images would be same for all methods in comparison.

Fig. 3 shows that HeNLM-3D method performs filtration better in fragments with small details (shown by ellipses in fig. 3, b and 3, c), and keeps some details which LjNLM-LR method blurs.

Conclusion

A three-dimensional method for CT image filtering is proposed. A test on real images shows that HeNLM-3D method performs filtering better than LjNLM-LR algorithm in areas that contain small details.

The research was supported by the RFBR grant 13-07-00584.



Fig. 2. The original CT image. A white border shows the fragment enlarged in fig. 3

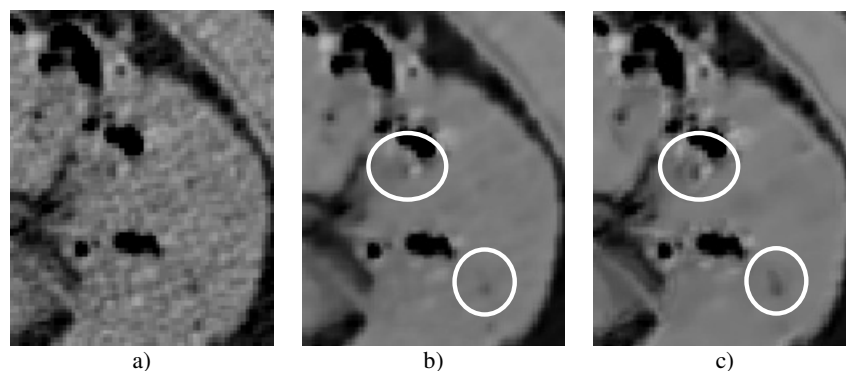


Fig. 3. The original CT image fragment (a) and the results of filtering by LjNLM-LR (b) and HeNLM-3D (c) algorithms

References

1. Natterer, F. *The Mathematics of Computerized Tomography* / F. Natterer. – SIAM : Society for Industrial and Applied Mathematics, 2001. – P. 222.
2. Can Noise Reduction Filters Improve Low-Radiation-Dose Chest CT Images? Pilot study / M.K. Kalra [et al.] // *Radiology*. – 2003. – Vol. 228. – P. 257–264.
3. 2.5D Extension of Neighborhood Filters for Noise Reduction in 3D Medical CT Images / M.V. Storozhilova [et al.] // *Lecture Notes in Computer Science*. – 2013. – Vol. 7870. – P. 1–16.
4. Kijewski, M.F. *The Noise Power Spectrum of CT Images* / M.F. Kijewski, P.F. Judy // *Phys. Med. Biol.* – 1987. – Vol. 32, no 5. – P. 565–575.
5. Buades, A. *A Non-Local Algorithm for Image Denoising* / A. Buades // *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. – 2005. – Vol. 2. – P. 60–65.
6. Manzanera, A. *Local Jet Based Similarity for NL-Means Filtering* / A. Manzanera // *20th Intern. Conference on Computer Vision and Pattern Recognition (ICPR)*. – 2010. – P. 2668–2671.
7. Gabor Feature Based Nonlocal Means Filter for Textured Image Denoising / S. Wang [et al.] // *Journal of Visual Communication and Image Representation*. – 2012. – Vol. 23. – P. 1008–1018.
8. Abramowitz, M. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. / M. Abramowitz, I.A. Stegun. – U.S. National Bureau of Standards, 1965.
9. Hermite Functions Expansion Based Non-Local Means Algorithm for CT Applications / N. Mamaev [et al.] // *Proc. of the 11th Intern. Conference “Pattern Recognition and Image Analysis: New Information Technologies”*. – 2013. – Vol. 2. – P. 638–641.
10. Knutsson, H. *Representing Local Structure Using Tensors* / H. Knutsson // *Proceedings 6th Scandinavian Conf. on Image Analysis*. – 1989. – P. 244–251.

VIDEO CONTENT ANALYSIS AND KEY FRAME EXTRACTION METHOD

S. Mashtalir, O. Mikhnova

Kharkiv National University of Radio Electronics, Ukraine

The latest advances in video summarization, namely key frame extraction, are examined in this paper. A new method based on Voronoi tessellations which provide more stable presentation of spatial frame content is described. Drawbacks and benefits of the method are highlighted along with its comparison with existing techniques. Future extensions are proposed that can enhance its application, quality and performance.

Introduction

Video refers to the most informative type of multimedia that unites graphical, sound and sometimes even text information. In addition, this information changes constantly in time, which is one of the main issues, as video processing speed plays an important role, and it decreases very much by increasing complexity of processing procedure. From another hand, too simple processing procedure does not guarantee the required quality. The main directions of video processing are the following: segmentation into objects and tracking, archiving and classification under genres, indexing and searching by request, annotation and summarization, scene boundary detection, identification of duplicates and commercials.

Lately, CBVR (Content Based Video Retrieval) issues have become urgent. So called semantic or high-level analysis is performed by artificial intelligence methods. However, while constructing CBVR systems, severe problems appear relating to a huge gap between traditional low-level features (used for image analyses for the latest decades) and high-level semantic presentation of video content. Two more problems arise from subjectivity of video perception by users and video volumes, the processing of which requires spatiotemporal segmentation under assumption of many properties and feature sets.

Quite a new direction of video processing, especially for our national science, is static video summarization, as a result of which users obtain stationary representative images (key frames) that characterize initial material and provide a brief overview of it. By analyzing existing foreign approaches to key frame extraction (table 1) [1–7], it is clearly seen that some groups of methods successfully extract frames with significant motion, others can only be applied to insignificant font changes, others are reasonable to be used for unchangeable lighting conditions, otherwise important frames can be missed or, on the contrary, key frames with high correlation between each other are extracted.

Table 1

Basic approaches to key frame extraction

Name of Approach	Representatives (Year of Publication)
Boundary-based approach	A. Nagasaka, Y. Tanaka (1991)
Matrix factorization	Y. Gong (2000), M. Cooper (2002)
Curve simplification	S. Lim (2001), K. Matsuda (2004), E. Bulut (2007)
Clustering	L. Li (2008), Z. Qu (2013)
Visual attention model	J. Peng (2010), L.J. Lai (2012), Q.-G. Ji (2013)
Genetic algorithms	X. Yang, Z. Wei (2011)
Histogram difference	B. Liang (2012), G.I. Rathod (2013)
Statistics	J. Almeida (2012), S.S. Kanade (2013)
Neural networks	D.P. Papadopoulos (2013)
Ontologies	K. Khurana (2013)

Thus, an urgent CBVR problem consists in studying of key frame extraction methods, finding the ways to increase quality and speeding-up extraction process under condition of procedure uniqueness for different types of videos. Construction of competitive methods for summarization requires searching for principally novel solutions, which is the goal of our research. This will give an ability not only to extract meaningful content from video, but also to eliminate information overload and guarantee user-friendliness and ease of access to stored video materials.

1. Spatiotemporal Segmentation for Key Frame Extraction

Video sequence is often segmented into scenes before key frames are extracted. Such a kind of temporal segmentation is called scene change identification, scene boundary detection, or SBD in short. Scene boundary

detection is usually performed by analysis of changes in color, texture, object shapes, motion or any other features selected for characterization of each frame in a video sequence. In our case it has been proposed to use scene boundary detection method introduced in [8]. It analyses time series constructed from object segmentation data, which provides an opportunity to take video content into account.

After any video has been segmented into scenes, key frames are selected by means of artificial intelligence methods and tools applied for graphical, structural, audio and textual information present in video. A sequence of key frames is obtained at the result, which forms a subset of initial video frames. It is the main phase of key frame extraction that has been introduced by authors of this article in [7].

The proposed method assumes using Voronoi tessellations for segmentation of frame content. Salient points are aided to construct Voronoi tessellations. These points are invariant to geometric and radiometric distortions under potential movement, they possess some common properties and reside at object borders or any other places that stand out against a background. Many other names can be met in literature to denote salient points: important points, meaningful points, interesting points, key points, characteristic points, sites, atoms and generating points [9–11]. Despite “generating points” can be met more often for Voronoi diagram construction, the term “salient” has been chosen to underline their significance for an image and rearrange them further for better content matching.

Harris method of salient point selection has been implemented for the purposes of current research. After initial placement of salient points they are rearranged by *k*-means clustering for better correspondence with image content to exclude local outliers of intensity, as this algorithm is considered to be good enough for searching of reasonable partitioning [12]. The applied *k*-means clustering algorithm assumes color and texture information. The algorithm has been chosen because of its correspondence with principles of Voronoi tessellations generated by means.

Fig. 1 shows examples of frame segmentation using Voronoi tessellations for news video file “factories_512kb.mp4” taken from Internet Archive (<http://archive.org>), lasting for more than 5 minutes and containing more than 50 scenes. From this figure it is clear that Voronoi tessellations look similar for nearly the same frames (consecutive frames with a man whose eyes look in front of him and then downwards, two latest frames with a little girl standing near a ruined building), and vice versa, tessellations look very different when frame content is not similar.



Fig. 1. Frames segmented by Voronoi tessellations

However, along with spatial changes in tessellations, meaningful features should be also taken into account. To compare Voronoi tessellations in this way, a special similarity metric has been introduced and implemented. Its uniqueness lies in assumption of selected feature set, which allows not only representing video content, but also stabilizing frame descriptions. For this purpose higher order Voronoi tessellations have been also applied [13].

Earlier researchers have not performed any comparisons of Voronoi tessellations. The exception is only Yukio Sadahiro, Japanese researcher who has compared Voronoi tessellations, though not with the

help of similarity metrics [13]. He has suggested using different methods of visual and quantitative analysis, including chi-square test, Cappa index and their extensions, area and perimeter of Voronoi tessellations, dispersion and standard deviation, centers of mass, etc. Sadahiro's contribution consists in comparison of different partition systems implemented for administrative region division in Japan. For this purpose he has used detailing density measure and hierarchical relations (such as inclusion, full and partial overlap).

However, it should be noted that area-based methods possess some duality from the point of video processing applications, as objects may be shot with different zoom. Different objects from video frames may have the same area, and vice versa, the same object in two frames may be of different area. Thus, video objects should not be tracked by considering their area-based properties only. In our case, different features should be included for frame comparison. Spatial features, textural and color features, motion estimations are among the main attributes used in CBIR and CBVR systems.

To extract key frames with lowest level of similarity, frames have been compared pairwise by taking color, texture and shape features into consideration. At the end of key frame extraction procedure similar or near-similar key frames are removed from the resulting sequence by comparison of all the key frames between each other using the proposed similarity metric. In other words, second pass of the algorithm is performed for the extracted frames. The second pass is much quicker than the first one, as less frames are left to be compared. It is used just in order to increase quality of finally extracted key frames by duplicate removal.

To test the performance and quality of key frame extraction procedure on various video genres, self-made full HD videos have been taken and open video test samples with lower resolution from Internet Archive, Open Video Project, Movie Content Analysis Project.

2. Benefits of the Proposed Procedure

Existing methods and artificial intelligence tools do not fully satisfy constantly increasing users' requirements for automatic video summarization. Extraction of meaningful video frames remains difficult without users' involvement. Moreover, a negative impact common for many methods is that the amount of key frames should be set a priori and frames are to be selected from video primarily segmented into scenes, which limits the amount of resultant key frames and sets restrictions without any assumption of content.

Calculations of precision and recall measures for the proposed and the existing methods have shown that the results obtained from curve simplification provide the highest precision, while the highest recall is provided by visual attention and histogram difference methods. In two latest cases too much frames are usually extracted, and additional measures should be provided to restrict this final amount. Genetic algorithms, neural networks and artificial immune networks, as well as motion analysis incorporated in visual attention model, are supposed to be "heavy" algorithms with additional extensions needed to perfect the resulting quality.

Boundary-based approaches possess the lowest values of both precision and recall as expected. Statistical approaches also cannot boast of outstanding results, though they are preferably used to ensure real-time processing. Ontologies assume too much users' involvement into key frame extraction or huge work on the underlying vocabulary and template dataset, which finally is applicable to a restricted branch of knowledge only.

The proposed method differs from existing ones in accuracy of obtained results and uniqueness of computational procedure which comprises of Voronoi tessellation comparison. Current algorithms provide opportunity to reveal changes in each frame, though these changes may be not that important from semantic point of view, while the proposed method enables to determine a set of key frames where only drastic changes are observed in content. Object borders may change greatly from frame to frame, but Voronoi tessellations are quite stable in these circumstances.

Fig. 2 shows linear diagram of Dice coefficients obtained under extraction of key frames from videos of different genres. Traditionally Dice coefficient unites such characteristics as precision and recall in a range from 0 to 1, which clearly demonstrates the estimations performed by respondents [14]. Twenty independent respondents have been attracted in our research.

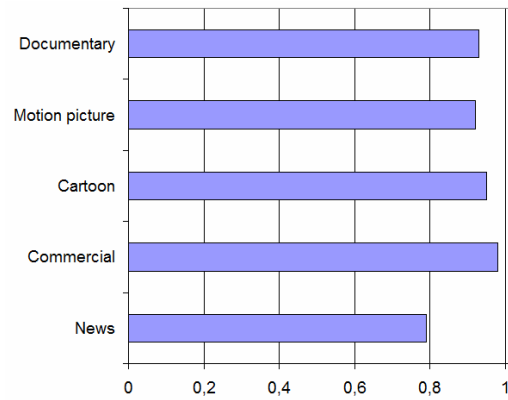


Fig. 2. Average values of Dice coefficient obtained after the estimation of key frames extracted from videos of different genres

Low level of Dice coefficient obtained for news videos can be explained by appearance of textual information in some of the test samples. Recognition of overlapping text, as well as incorporation of text mining and processing methods, was not the goal of our research.

Fig. 3 illustrates average time needed for processing of each kind of video lasting for a minute. Faster processing of documentaries and commercials has been performed because of existence of several black and white samples in these two test collection categories. The less detail a frame contains, the less salient points are detected and the faster processing speed is. Initial identification of salient points and their rearrangement by *k*-means clustering takes the longest time in the procedure. If there were no black and white samples in the collections, it would take approximately 2 minutes to process each frame.

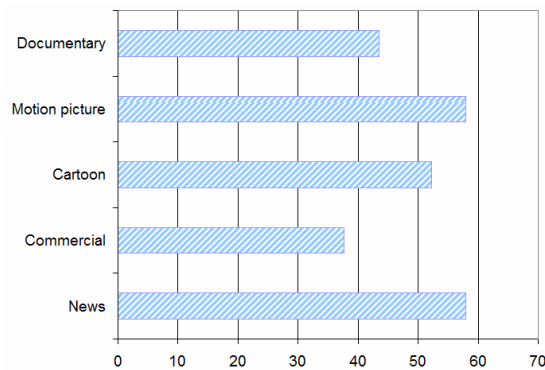


Fig. 3. Average time (in minutes) needed for processing of each kind of video material lasting for a minute

Such a huge amount of processing time is obtained because of several reasons. First of all, each of the selected tracks contains 29 frames per second. Secondly, all the data have been processed at a low frequency of 1.6 GHz Dual CPU with only 1 GB RAM. Along with numerous content details that decrease processing speed, 5 salient points are built and refined for a second, whereas 15 salient points for the same frame are constructed for 2 seconds. This is the case for 360*240 pixels, but the processing time increases for greater frame sizes. Unfortunately, video processing is performed far from a real-time mode, though our prime goal was the quality.

By comparing our current results with the results received after key frame extraction by some other methods, one may come to the conclusion that our method can take into account significant changes in content and omit slight changes due to matching of Voronoi mosaics. Upon this, there is no need to set the amount of key frames a priori, as it has been done for some clustering approaches. Our method can extract frames from homogeneous scenes as well as heterogeneous scenes with significant motion and background changes, which is very important from the point of procedure uniqueness. As all the rest of the methods, ours is also sensitive to resolution. Concerning processing speed which is very low in our method, it should be noted that many other methods also do not aim at real-time processing, but the quality only that is too hard to reach by now.

3. Conclusions and Future Work

Analysis of existing key frame extraction techniques has been performed. It has been shown that the main problems for video processing (and key frame extraction in particular) are complexity of segmentation into objects under conditions of non-stationary background and object overlap, traditional conflict between low- and high-level content presentation and subjectivity of result estimation.

The necessity of a new method development (that could process unstable content of different genres with an equal efficiency) is evident. The proposed method is based on comparison of Voronoi tessellations that are constructed for each frame. It also takes scene boundaries into account, reduces duplicates at the resulting key frame sequence, and it is totally competitive with existing methods in terms of quality (not in time), which has been testified by comparison of precision and recall measures for the proposed and the existing methods using the same open source video samples.

Application of Voronoi diagrams gives a chance to increase stability of partitions, and consequently key frame extraction robustness due to consideration of regions around salient points. Salient points themselves provide less information than the whole regions, moreover, their shifts are more significant from frame to frame. Also it is hard to distinguish a point that is moved, from a point that is left in its old position. Voronoi diagrams are more preferable from the point that there is no need in video segmentation into real objects, as there are still hardships to detect objects in videos of different genres with dynamically changing font.

In future the proposed method can be extended by weighted feature estimates, as in our case a linear combination of features has been used. Along with color, texture and shape features, additional features can be applied. Of course, attempts towards near-real-time mode can be performed. It can be reached, first of all, by incorporating FAST (Features from Accelerated Segment Test) detector instead of Harris algorithm, avoiding long lasting computation of derivatives. Instead of k -means clustering, Simulated Annealing in combination with Tabu Search, introduced in [15], can be also implemented to optimize key frame extraction. Highest order Voronoi diagrams can be applied at the first step of key frame extraction to identify scene boundaries and not to overload computations by supplementary methods. Additionally, not very high frame size and resolution can be used (in case there is such a choice, nearly 800*600 is enough) to decrease the number of details, and thus salient points detected, in each frame.

References

1. Lai, L.J. Key frame extraction based on visual attention model / L.J. Lai, Y. Yi // Journal of Visual Communication and Image Representation. – 2012. – Vol. 23, no. 1. – P. 114–125.
2. Rathod, G.I. An Algorithm for Shot Boundary Detection and Key Frame Extraction Using Histogram Difference / G.I. Rathod, D.A. Nikam // International Journal of Emerging Technology and Advanced Engineering. – 2013. – Vol. 3, no. 8. – P. 155–163.
3. Almeida, J. VISON : VIdeo Summarization for ONline applications / J. Almeida, N.J. Leite, R.S. Torres // Pattern Recognition Letters. – 2012. – Vol. 33, no. 4. – P. 397–409.
4. Kanade, S.S. Dominant color based extraction of key frames for sports video summarization / S.S. Kanade, P.M. Patil // International Journal of Advances in Engineering & Technology. – 2013. – Vol. 6, no. 1. – P. 504–512.
5. Khurana, K. Video Annotation Methodology Based on Ontology for Transportation Domain / K. Khurana, M.B. Chandak // International Journal of Advanced Research in Computer Science and Software Engineering. – 2013. – Vol. 3, no. 6. – P. 540–548.
6. Automatic summarization and annotation of videos with lack of metadata information / D.P. Papadopoulos [et al.] // Expert Systems with Applications. – 2013. – Vol. 40, no. 14. – P. 5765–5778.
7. Mashtalir, S.V. Stabilization of key frame descriptions with higher order Voronoi diagram / S.V. Mashtalir, O.D. Mikhnova // Bionics of Intelligence. – 2013. – No. 1. – P. 68–72.
8. On-line video segmentation using methods of fault detection in multidimensional time sequences / Y. Bodyanskiy [et al.] // International Journal of Electronic Commerce Studies. – 2013. – Vol. 3, no. 1. – P. 1–20.
9. Bradski, G. Learning Open CV. Computer Vision with the OpenCV Library / G. Bradski, A. Kaehler. – O'Reilly Media, 2008. – 578 p.
10. Spatial tessellations: Concepts and applications of Voronoi diagrams / A. Okabe [et al.]. – 2nd ed. – Wiley, 2000. – 671 p.

11. Aurenhammer, F. Voronoi Diagrams – A Survey of a Fundamental Geometric Data Structure / F. Aurenhammer // ACM Computing Surveys. – 1991. – Vol. 23, no. 3. – P. 345–405.
12. Mashtalir, S.V. Key frame recognition using Voronoi tessellations. Visnyk of National University “Lviv Polytechnic” / S.V. Mashtalir, O.D. Mikhnova // Computer Science and Information Technologies. – 2013. – Vol. 751. – P. 52–58.
13. Sadahiro, Y. Analysis of the relationship among spatial tessellations / Y. Sadahiro // Journal of Geographical Systems. – 2011. – Vol. 13, no. 4. – P. 373–391.
14. Manning, C.D. Introduction to Information Retrieval / C.D. Manning, P. Raghavan, H. Schütze. – Cambridge University Press, 2008. – 496 p.
15. Chu, S.-C. Improved Clustering and Soft Computing Algorithms / S.-C. Chu // Doctoral Thesis. – Adelaide, 2004. – 238 p.

FAST IMAGE MATCHING WITH SLIDING HOGS

D. Miramontes-Jaramillo¹, V. Kober^{1,2}, V. Karnaukhov²

¹Department of Computer Science, CICESE, Ensenada, México;

²Institute for Information Transmission Problems of RAS, Moscow;

e-mail: dmiramon@cicese.edu.mx, vkober@cicese.mx

Various vision applications utilize matching algorithms for searching a target object in a scene image. We present a fast matching algorithm based on recursive calculation of histograms of oriented gradients (HOGs). In order to speed up the algorithm we propose pyramidal image decomposition and parallel implementation with modern multicore processors. The algorithm yields a good invariance performance for both in-plane and out-of-plane rotations of a slightly scaled scene image. Computer results obtained with the algorithm are compared with those of common algorithms.

Introduction

Recently, numerous matching algorithms using features or keypoints were proposed. Among them, Scale Invariant Feature Transform (SIFT) [1] and Speeded-Up Robust Features (SURF) [2] are widely used algorithms. Although feature-based matching methods are popular, template-based matching algorithms are an attractive alternative for real-time applications, the later extract features as statistical moments, color and intensity histograms [3, 4]. Another approach is a combination of feature-based and template matching algorithms. For instance, the Scale Invariant Compressed Histogram Transform (SICHT) [5] uses, as features, the histograms of oriented gradients [6] calculated in a moving window.

The proposed algorithm performs the matching of histograms of oriented gradient in several circular windows running across an input image, and final decision is taken on the base of joint matching results for all the windows; for speed up, the algorithm utilizes decimation, two histogram resolutions, and local threshold filtering of histograms; for noise removal an adaptive neighborhood process is used; and there exists a trade-off between complexity and accuracy of the algorithm controlled by desirable values of the probabilities of false alarms and miss errors, and processing time requirement.

The performance of the proposed algorithm in a test database is compared with that of the SIFT, the SURF and a new efficient Oriented FAST and Rotated BRIEF (ORB) algorithm [7] in terms of matching accuracy and processing time.

1. Proposed Algorithm

First, let us define a set of circular windows $\{W_i = 1, \dots, M\}$ in a reference image as a set of the following closed disks:

$$W_i = \{(x, y) \in \mathbb{R}^2: (x - x_i)^2 + (y - y_i)^2 \leq r_i^2\}, \quad (1)$$

where (x_i, y_i) are coordinates of the window center and r_i is the radius of the i^{th} disk. The disks form a geometric structure with relative distances and angles between the window centers.

The structure runs across a scene image. The search of the geometric structure in the scene image helps us to reduce false alarms and miss errors as the joint probability of occurrence of several events is less than that of a single event. The histograms of oriented gradients are calculated in circular areas and used for matching. It is interesting to note that at any position of the structure each disk contains image area that is unchangeable during rotation; therefore, the histogram of oriented gradients computed in a circular window is also invariant to rotation. The use of a single window for matching of the histograms of oriented gradients yields low matching results. It is recommendable to choose a minimum number of equal disks to fill inside as much as possible the reference object. Each circular window contains only information of the object without background data that can modify the histograms of oriented gradients.

Actually, numerous experiments showed that the number M of circular windows may be chosen from 2 to 4 to yield the matching performance comparable with that of the SIFT. Note that the processing of the structures is done only from the reference image once.

Histograms of oriented gradients are good features for matching [6] because they possess a good discriminant capability and are robust to small image deformations such as rotation and scaling. First, at each position of the i^{th} circular window on a scene image we compute gradients inside the window with

the help of the Sobel operator [8]. Next, using the gradient magnitudes $\{Mag_i(x, y) : (x, y) \in W_i\}$ and orientation values quantized for Q levels $\{\varphi_i(x, y) : (x, y) \in W_i\}$, the histogram of oriented gradients can be computed as follows:

$$HoG_i(x) = \begin{cases} \sum_{(x,y) \in W_i} \delta(\alpha - \varphi_i(x, y)), & Mag_i(x, y) \geq Med \\ 0, & otherwise, \end{cases} \quad (2)$$

where $\alpha = \{0, \dots, Q - 1\}$ are histogram values (bins), Med is the median of the gradient magnitudes inside of the circular window, and δ is the Kronecker delta function. Note that the computation in Eq.(2) requires approximately $\lceil \pi r_i^2 \rceil$ addition operations, where $\lceil \cdot \rceil$ is an integer part operator. In order to reduce computational complexity the calculation of the histograms at any position of the sliding window can be performed in a recursive manner as follows:

$$HoG_i^k(x) = HoG_i^{k-1}(x) - \sum_{(x,y) \in OutP_i^{k-1}} \delta(\alpha - Out\varphi_i^{k-1}(x, y)) + \sum_{(x,y) \in InP_i^k} \delta(\alpha - In\varphi_i^k(x, y)) \quad (3)$$

where $OutP_i^{k-1}$ is a set of outgoing orientation values whose pixels belong to the half of the perimeter of the sliding window at step $k - 1$, that is,

$$(4)$$

InP_i^k is a set of incoming orientation values whose pixels belong to the half of the perimeter of the sliding window at step k given by

$$(5)$$

The computational complexity of this calculation is approximately πr_i^2 addition operations. In addition, the median computation does not limit the power of the method; a magnitude histogram is computed alongside the orientation histogram and a pointer of the median position at step $k - 1$ is stored; at step k before the orientation histogram is updated, we update the median value with the incoming and outgoing magnitude values in the magnitude histogram as well as its position, doing so also takes as we only modify that fraction of the histogram. The proposed method does not use a linear low-pass filtering like common algorithms do, because such filtering could completely destroy or attenuate gradients. The algorithm further rejects a noise influence by means of local adaptive threshold filtering simultaneously with the histogram update.

To provide rotation invariance we utilize a normalized correlation operation for comparison of the histograms of reference and scene images. Let us compute a centered and normalized histogram of oriented gradients of the reference image as follows:

$$\overline{HoG_i^R}(\alpha) = \frac{HoG_i^R(\alpha) - Mean^R}{\sqrt{Var^R}}, \quad (6)$$

where $Mean^R$ and Var^R are sample mean and variance of the histogram, respectively.

The correlation output for the i^{th} circular window at the position k can be computed with the help of the Inverse Fast Fourier Transform [8],

$$C_i^k(\omega) = IFT \left[\frac{HS_i^k(\omega)HR_i^k(\omega)}{\sqrt{Q \sum_{q=0}^{Q-1} (H \circ G_i^k(q))^2 - (HS_i^k(\omega))^2}} \right] \quad (7)$$

where $HS_i^k(\omega)$ is the Fourier Transforms of the histogram of oriented gradients inside of the i^{th} circular window over the scene image and $HR_i^k(\omega)$ is the Fourier Transform of $\overline{H \circ G_i^k(\omega)}$; the asterisk denotes complex conjugate. The correlation peak is a measure of similarity of two histograms. The correlation peaks are in the range of $[-1, 1]$. It is of interest to note that the normalized correlation peaks possess two important properties: first, invariance to rotation because a cyclic shift of the histogram values corresponds to a cyclic shift of the correlation output and does not change the correlation peak value; second, the normalization in eq.(6) and eq.(7) helps us to take into account a slight scale difference between the reference and scene images. Computation of the centered and normalized histograms for all circular windows over the reference image as well as the Fourier Transforms can be done as offline preprocessing.

For speed-up of the matching we do not use the classical pyramidal approach, which is based on a low-pass filtering and subsequent subsampling the resultant image to get a lower resolution image of a smaller size [9]. A simple decimation along rows and columns of an input image is more suitable to accelerate the matching of histograms of oriented gradients. We perform the decimation of the original image with a step of L , that is, L^2 small images composed by pixels of the original image are created.

Finally, the matching is carried out between the first small scene image and all small reference images. Remind that the decomposition of a reference image into a set of small images can be done as preprocessing. A similar technique is popular for estimation of motion vectors and for super resolution image reconstruction [10, 11], but until our knowledge it has not been applied to image matching. Another important parameter that can accelerate the matching, is the number of histogram bins Q . We suggest a two-pass procedure: first, perform independent fast matching for all circular windows with a reduced number of bins, say Q_R and a decision threshold. The objective is to reduce as much as possible the number of points to be considered in the second matching pass keeping a low value of the probability of miss errors. Second, only accepted points are considered to carry out the matching with a regular number of bins Q for each window and a decision threshold T_{H_Q} . Finally, analysis of joint appearance of all assigned circular areas of the reference is able to further reduce the probability of miss errors as well as the probability of false alarms.

Take final decision about the presence of the reference object at the position k . During the analysis we consider possible deformation of the original structure owing to scaling and out-of plane rotation. Start the analysis at the position of the highest peak for the first window. Next, assign a ring area with the internal and external radii for searching the peak inside the second window and so on. The search procedure continues either until the correlation peaks of all circular windows will be found in the right areas (good matching) or for any circular window the correlation peak will not be found in the predicted area (failed matching). If the reference object was detected at the position k , then the next search area to be considered is separated from the current position by the size of the reference object.

2. Experimental Results

In this section we present experimental results using the image database ALOI [12]. The database is composed of 100 objects with wide range of out-of-plane rotations and illumination schemes; it features each image in a set of resolutions and image formats as well as a mask for every object image. Ten scene images with the size of 1280×1024 pixels and ten reference images of the size 144×144 pixels with different objects were used. Each reference object is located randomly inside the scene images in 10 different positions.

The performance of the proposed algorithm was compared with that of the SIFT, SURF and efficient algorithm referred to as ORB. The parameters of the algorithms were taken from corresponding papers [1, 2, 7]. The algorithms are tested in different conditions such as in-plane/out-of-plane rotations and a slight scaling. The performance is evaluated in terms of the number of correct matches and processing time.

The proposed algorithm referred to as CWMA uses two circular windows for each object with a radius r (dependent on the size of object). For a better matching we use $Q_R = 16$ and $Q = 64$ bins instead of bins proposed in [6]. The parameters of the algorithm are as follows: $M = 2$, $Q = 64$, $Q_R = 16$, $L = 4$. The decision thresholds are assigned as follows: for $Q_R = 16$ (first pass) the threshold $Th_{QR} = 0.8$ yields the probability of miss errors on a test database $P_{ME} = 0.05$ and accepts about of points to be considered on the second pass with $Q = 64$; for $Q = 64$ (second pass) the threshold $Th_Q = 0.7$ provides the probability of miss errors of 0.05 and the probability of false alarms of 0.21 . The final decision reduces the probability of false alarms to 0.05 .

The performance of the tested algorithms is resumed in the table, where one can see the average values of Hit/Miss rate as well as the average processing time for the tested algorithms.

Comparison of tested algorithms in terms of Hit/Miss averagerate and processing time

Algorithm	Rotation in-plane	Rotation out-of-plane	Scale	Time
CWMA	98.25%	77.62%	94.11%	1.7 sec.
SIFT	98.36%	66.52%	95%	9.82 sec.
SURF	72.65%	51.46%	74.44%	0.95 sec.
ORB	85.75%	61%	84.55%	1.7 sec.

Conclusion

This paper presented a fast image matching algorithm based on recursive calculation of oriented gradient histograms over several circular sliding windows was presented. Multicore processors with inherent parallel architectures can help to implement the algorithm for image matching with large scenes at high rate. Experimental results showed that the proposed algorithm outperforms the common algorithms for in-plane rotation, yields a similar performance to that of the SIFT for out-of-plane rotation, and requires processing time close to the SURF. The proposed algorithm is attractive for real-time applications such as tracking when rotation invariance matching with a slight scaling is required.

References

1. Lowe, D.G. Object recognition from local scale-invariant features / D.G. Lowe // Proc. Intern. Conf. ComputerVision. – 1999. – Vol. 2. – P. 1150–1157.
2. SURF: Speeded Up Robust Features / H. Bay [et al.] // Comput. Vis. Image Underst. – 2008. – Vol. 110 (3). – P. 346–359.
3. Manzurv, T. Optical correlator based target detection, recognition, classification, and tracking / T. Manzurv, J. Zeller, S. Serati // Applied Optics. – 2012. – Vol. 51. – P. 4976–4983.
4. Tsia, D.M. Rotation-invariant pattern matching using color ring projection / D.M. Tsia, Y.H Tsai // Pattern Recognition. – 2002. – Vol. 35 (1). – P. 131–141.
5. Zalesky, B.A. Scale Invariant Algorithm to Match Regions on Aero or Satellite Images / B.A. Zalesky, P.V. Lukashevich // Proc. Pattern Recognition and Information Processing. – 2011. – Vol. 11. – P. 25–30.
6. Dalal, N. Histograms of Oriented Gradients for Human Detection / N. Dalal, B. Triggs // Computer Vision and Pattern Recognition. – 2005. – Vol. 1. – P. 886–893.
7. ORB: an efficient alternative to SIFT or SURF / H. Rublee [et al.] // IEEE Intern. Conf. on Computer Vision. – 2011. – P. 2564–2571.
8. Pratt, W.K. Digital Image Processing / W.K. Pratt. – John Wiley & Sons, 2007.
9. Lindeberg, T. Scale-space theory: A basic tool for analysis structures at different scales / T. Lindeberg // Journal of Applied Statistics. – 1994. – Vol. 21 (2). – P. 225–270.
10. López-Martínez, J.L. Fast image restoration algorithm based on camera microscanning / J.L. López-Martínez, V. Kober // Proc. SPIE 7443. – 2009. – P. 744310–744315.
11. Liu, B. New fast algorithms for the estimation of block motion vectors / B. Liu, A. Zaccarin // IEEE Transactions on Circuits and Systems for Video Technology. – 1993. – Vol. 3 (2). – P. 148–157.
12. Geusebroek, J.M. The Amsterdam library of object images / J.M. Geusebroek, G.J. Burghouts, A.W.M. Smeulders // Intern. J. Computer Vision. – 2005. – Vol. 61 (1). – P. 103–112.

REDUCING ENERGY CONSUMPTION IN WIRELESS SENSOR NETWORKS BY CICA ALGORITHM IN COMPARISON WITH THE GENETIC ALGORITHM

R. Askari Moghaddam¹, S.M. Mortazavi², M. Malekzadeh³

¹ Faculty of New Sciences and Technologies, University of Tehran, Iran;

² Master of Science in Software Engineering, Payam Noor University, Tehran, Iran;

³ Faculty of Engineering, Golestan University, Gorgan, Iran

e-mail: r.askari@ut.ac.ir

Wireless sensor network is a collection of small sensor node in which the ability to sense its environment and monitoring and sending data to a base station. One of the main challenges in sensor networks is energy constraint of the nodes which directly affects the network lifetime. Hence, there are various methods to optimize the power consumption which in turn increase the lifetime of the sensor networks. In this paper, a new method called CICA⁴ for clustering wireless sensor networks using Imperialist Competitive algorithm (ICA) is presented that divisions sensor nodes to the balanced clusters. Result of its simulation shows the successful performance to increase the lifetime of wireless sensor networks rather than the HCR algorithm that is based on genetic algorithm.

Introduction

In recent years, the telecommunications and electronic technology advances causes the construction of the small and cheap sensors that connected with each other via a wireless network [1]. These networks that called wireless sensor networks have become as a tool for extracting data from the environment and monitor environmental events and their applications in the fields of domestic, industrial and military increasing day-to-day [2].

Each node in this network consists of three main part included sensor, processor and wireless transmitter/receiver and the data received from the environment such as temperature, humidity, pressure, light, motion and etc sent for processing to the base station. Constraint of the power supply in this network is a basic challenge, therefore offering energy efficient methods that increase life of nodes and network is interesting for researchers. In this paper, a new method for clustering wireless sensor networks using Imperialist Competitive algorithm (ICA) is presented.

1. Related Works

1.1. Clustering Nodes

The main task of the sensor network nodes is collect information from the environment. one of the main reasons for the energy consumption of the nodes is the data transferring and consuming energy by sending Information wirelessly is directly related to the second (or higher) exponent of distance between source and destination in the network. Therefore plans that will shorter communication distance between the nodes can reduce the energy consumption and increase the lifetime of sensor networks [3].

The ideal mode of the sensor network energy consumption is that the energy of all the nodes finish together, thus to increase the lifetime of the network we try network load distribution be in a uniform distribution until time interval between the death of the first node and the death of the last node be at least. To achieve this goal, several communication protocols have been proposed. So far protocols based on clustering considerably slows down energy consumption the networks [4].

Clustering is that the entire network partitioning into several independent groups that called cluster and each node located in one of that clusters and one of the nodes in each cluster node selected as a cluster head.

The duty of all cluster heads is collecting data from the nodes in its cluster and then sending this information directly or with other cluster heads to the base station. Thus, clustering can greatly reduce communications costs.

⁴ Clustering with Imperialist Competitive Algorithm

1.2. HCR⁵ algorithm:

Hierarchical cluster-based routing algorithm (HCR) [5] is an extension of the LEACH⁶ protocol [6] that is a self organized cluster-based approach for continuous monitoring. In LEACH, the network is randomly divided into several clusters, where each cluster is managed by a cluster head (CH). The sensor nodes transmit data to their cluster heads, which transmit the aggregated data to the base station. In HCR, each cluster is managed by a set of associates and the energy efficient clusters are retained for a longer period of time; the energy-efficient clusters are identified using heuristics-based approach. Moreover, in a variation of HCR, the base station determines the cluster formation. A Genetic Algorithm (GA) is used to generate energy-efficient hierarchical clusters. The base station broadcasts the GA-based clusters configuration, which is received by the sensor nodes and the network is configured accordingly. For continuous monitoring applications, the simulation results show that HCR is more energy efficient than the traditional cluster-based routing techniques.

1.3. Imperialist Competitive Algorithm

So far variety of evolutionary algorithms is provided for optimization such as Genetic Algorithm (GA), Particle Swarm Optimization (PSO), Simulated Annealing (SA) and other existing algorithms in this field. Recently, a new algorithm called Imperialist Competitive Algorithm is presented by Atashpaz-Gargari and Lucas in 2007 [7] that based on humans' political-social developments.

Like other evolutionary ones, the Imperialist Competitive algorithm starts with an initial population which is called country. A number of countries, best in population, are selected as imperialist and the others are their colonies.

2. CICA Algorithm

Here we review the CICA algorithm which is a clustering algorithm in wireless sensor networks and based on Imperialist Competitive Algorithm. Due to Proof of optimality, we compare CICA algorithm with the HCR algorithm in wireless sensor network clustering that is based on genetic algorithm and we show the simulation results with the same parameters in both algorithm.

Effective parameters in CICA Algorithm are:

- distance between Sensor nodes and the cluster heads: The main parameters considered in the CICA algorithm is sum of distance between all nodes of their cluster head. If the cluster heads be selected so that this number is smaller, then we will more efficient clustering;

- distance of nodes to the base station: The nodes that their distance to the base station is less than their distance from their cluster head, instead sending data to their cluster head, sending data directly to the base station for avoid additional consumption of energy;

- the remaining energy of the sensor nodes: when energy of a cluster head finished, selecting cluster head is done for the cluster nodes on which the node is selected as a cluster head that among nodes of cluster have the highest residual energy.

3. Results of Simulation

Here we review the results of the simulation. The simulation was done in MATLAB environment and simulation results performed in three parts consist of compare the number of alive nodes, the mean energy consumption of the nodes and compare the total energy consumption of all nodes have been shown. Initial values of the common parameters of wireless sensor network simulation is presented (table 1). These parameters for both CICA algorithm and HCR algorithm are same.

Table 1
Common parameters values in simulations scenarios

Common parameters values
Initial Energy = 0.1 j
ETX = 50*0.000000001 (j/bit)

⁵ Hierarchical Cluster-based Routing

⁶ Low Energy Adaptive Clustering Hierarchy

ERX = 50*0.000000001 (j/bit)
Efs = 10*0.000000000001 (j/bit/m2)
Eamp = 0.0013*0.000000000001 (j/bit/m4)
EDA = 5*0.000000001 (j/bit/signal)
do = 87.7 (m)
Packet Size = 4000 (bit)
Network Size = 100 * 100 (m)
Number Of Nodes = 100
Base Station Position = (50 , 50)

3.1. Number of alive nodes

By finishing energy of a node, that node dies. One of the main parameters for algorithm optimality in wireless sensor networks is time of first node die, and accordingly, time of other nodes death. To display the number of nodes alive parameter in the sensor network, we calculated the total number of nodes in the current round that have energy.

In (fig. 1) shows the number of alive nodes per round in CICA algorithm and HCR algorithm. CICA algorithm first node dies occurs at round 197 and the first node dies in the HCR algorithm at round 164.

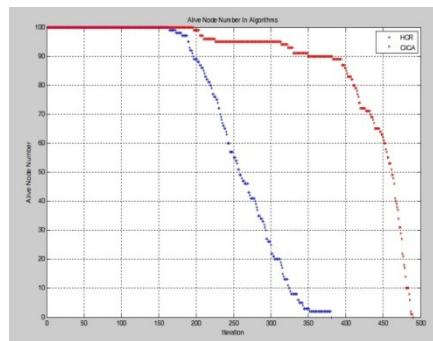


Fig. 1. Compare of the number of alive nodes

So CICA algorithm delay first node die rather than HCR algorithm. The analysis of (fig. 1) observed that the HCR algorithm has lost more than 70 percent of nodes at round 293 while in the CICA algorithm it happens at round 474, that is a reason for optimality of CICA algorithm in delay death of nodes and thus increase the efficiency of wireless sensor networks.

3.2. Compare average energy consumption per node

In (fig. 2) the parameter of average energy consumption per node until the first node dies in the HCR algorithm and CICA algorithm are compared. Average energy consumption per node until the first sensor node dies is equal to the total energy consumed until the death of the first node divided by the total number of nodes in the sensor network.

In this comparative parameter which algorithm is more efficient that have lower of average energy consumption per node that causes the life of the sensor networks be larger that is our purpose.

Since the denominator of the equation means the total number of sensor nodes is the same for both algorithms (in this case 100), which algorithm is more optimal that have smaller numerator that means the total energy consumed by the algorithm until the death of the first node is less than the other algorithm.

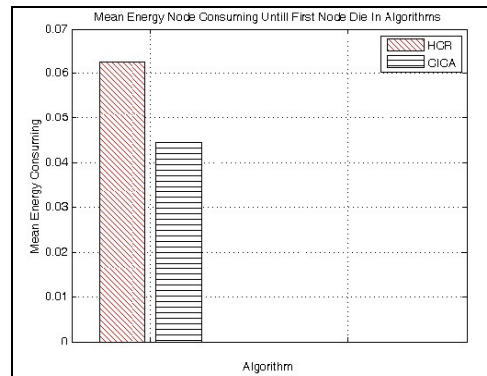


Fig. 2. Compare average energy consumption per node

The analysis of (fig. 2) is observed that CICA algorithm has an average 0.0445 j for each network node until the first node death which is lower (more optimal) and the HCR algorithm has an average consumption per node 0.0625 j.

This optimized average of energy consumption because of balanced clusters in wireless sensor network and thus consumption is balanced by CICA algorithm that is a major advantage in the network.

3.3. Compare of the total energy consumption of all nodes

In the (fig. 3) we compared the total energy consumption of all the nodes in each round of the CICA algorithm and HCR algorithm. In this comparative parameter which algorithm is more efficient that have less of the total energy consumption of all nodes. The analysis of (fig. 3) is observed that the CICA algorithm is more optimal than HCR algorithm in all rounds.

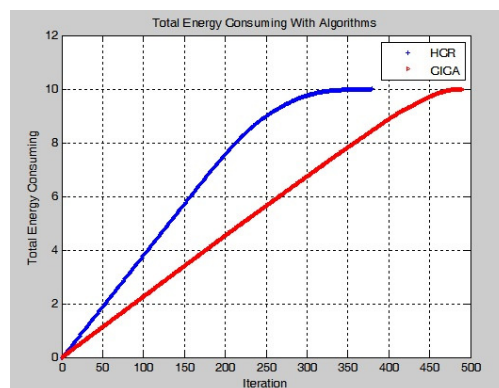


Fig. 3. Compare total energy consumption of all nodes

For example in the round 100, the total energy consumption of all nodes in the CICA algorithm is 2.2495 j and for the HCR algorithm is 3.7950 j. reducing the total energy consumption in CICA algorithm causes of finding optimal place for cluster heads. Because of computing the optimum location of cluster heads for a minimum distance of nodes of their cluster heads, nodes consume less energy in cluster heads to sending data.

Moreover, according to (fig. 3) is observed at round 200 of algorithms, the HCR algorithm more than 75 % of total energy consumed, but in CICA algorithm more than half of its total energy remains.

Conclusion

One of the main challenges in sensor networks is energy constraint of the nodes which directly affects the network lifetime. In this paper, a new method called CICA for clustering wireless sensor networks using Imperialist Competitive algorithm is presented that divisions sensor nodes to the balanced clusters. Result of its simulation shows the successful performance to increase the lifetime of wireless sensor networks. Benefits of CICA algorithm for clustering sensor nodes in a wireless sensor network comparable to the HCR algorithm is the following:

– creating balanced clusters: Intended to measure the distance from the specified cluster heads in the CICA algorithm causes the homogeneous and balance clustering thus solve one of the biggest disadvantages in the HCR algorithm that is the lack of a balanced clusters and the cluster heads near each other;

– selecting nodes with more energy as a cluster heads: Since the energy consumption of cluster heads is more than normal nodes in wireless sensor networks, selecting cluster heads with more residual energy among nodes in a cluster in CICA algorithm causes more life time of network.

Ultimately, in this paper we provide an evolutionary clustering method called CICA for energy optimization in wireless sensor networks with the successful results in its simulation.

References

1. Heinzelman, W. An Application-Specific Protocol Architecture for Wireless Microsensor Networks / W. Heinzelman, A. Chandrakasan, H. Balakrishnan // IEEE Transactions on Wireless Communication. –2002. – Vol. 1. – P. 660–670.
2. A Survey on Sensor Networks / I. Akyildiz [et al.] // IEEE Communications Magazine. – 2002. – Vol. 40. – P. 102–114.
3. Zhang, Y. RFID AND Sensor Networks / Y. Zhang, T. Yang, J. Chen // AUERBACH Pub. – CRC Press, 2010. – Ch. 2.
4. Younis, O. Node Clustering in Wireless Sensor Networks : Recent Developments and Deployment Challenges. / O. Younis, M. Krunz, S. Ramasubramanian // IEEE Network (special issue on wireless sensor networking). – 2006. – Vol. 20. – P. 20–25.
5. Hussain, S. Genetic algorithm for hierarchical wireless sensor networks / S. Hussain, A.W. Matin, O. Islam // Journal of Networks. – 2007. – P. 87–97.
6. Heinzelman, W. Energy Efficient Communication Protocol For Wireless Microsensor Network (Leach) / W. Heinzelman, A. Chandarakasan, H. Balakrishnan // Proc. of 33rd Hawaii International Conference Systems Science. – 2004. – Vol. 8. – P. 3005–3014.
7. Atashpaz-Gargari, E. Imperialist Competitive Algorithm: An Algorithm for Optimization Inspired by Imperialist Competition / E. Atashpaz-Gargari, C. Lucas // IEEE Congressman Evolutionary Computation. – 2007. – P. 4661–4667.

APPLICATION OF STRUCTURE TENSOR FEATURES FOR DESCRIBING AN ARTISTIC MANNER

D. Murashov

Dorodnicyn Computing Centre of RAS, Moscow

e-mail: d_murashov@mail.ru

In this paper, a task of comparing images for the purposes of attribution of paintings is considered. A feature description for comparing textural fragments characterizing the artistic style of a painter is proposed. Selected image features give quantitative description of artistic style and provide suitable accuracy of features computing. Results of the computing experiment showed the efficiency of the proposed features. Proposed feature set may be used as a part of technological description of fine art paintings for restoration and attribution.

Introduction

A great interest to applying computer techniques to the study of fine arts is shown during last decades [1]. One of the trends in attribution today is related to the analysis of digital images of paintings and called as "Computer-assisted Connoisseurship" [2]. In the glossary of the National Gallery [3], attribution of paintings is defined as an assessment of who was responsible for creating a particular work. Sometimes the term "attribution" is interpreted more widely and includes also an assessment of art school, time, country, etc. [4].

In this paper, a task of comparing images of fine art paintings for the purposes of attribution is considered. The idea of applying image analysis techniques for attribution is that to compare images of authentic and studied paintings by features characterizing individuality of an artist.

In publications of different research groups the following approaches to the task have been proposed. The first is based on the exhaustive comparison of square image fragments of the researched images [5]. The features as a rule are derived from the coefficients of the orthogonal transforms (in particular, wavelet transform). This approach is of high computational complexity and is sensitive to conditions of image acquisition and hardware parameters [6]. The second approach provides features computed from the segmented brushstrokes [7–10]. But usually it can be found too few paintings with a sufficient number of distinguishable brushstrokes that can be successfully segmented in automatic mode, or even manually. It can be done, for example, in images of paintings by van Gogh, P.J. Pollock, and some others. It is, therefore preferable to use features that can be computed directly from images bypassing brushstroke segmentation.

In this work the images of portraits are analyzed. Individuality of an artist the experts associate with brushwork features. In accordance to recommendations of art experts [4], we propose to use for comparing pictures a group of brushstrokes that form any detail of a picture, or the boundary between details and background. For example, in the portraits such details are lips, chin, nose, forehead, eyes, folds of clothes, etc. In [7] for attribution of portrait miniatures, the homotypic fragments of the human faces were compared. The fragments were segmented using geometric model of face developed by the authors. In this work we also use images of the homotypic informative face details: forehead, nose, and cheek. It should be noticed that the size of the images in current research is much larger than in [7], and selected image fragments differ from those in [7]. The fragments of three types are shown in fig. 1.

The problem is formulated as follows. Let U_j , $j = 1, 2, \dots, J$ be images of paintings by J authors; $U : R^2 \times R \rightarrow R$. Let an informative fragment u_j^i of type i taken from image U_j is characterized by a feature $x_j^i = \gamma(u_j^i)$, $i = 1, 2, \dots, I$, $\gamma : R^2 \times R \rightarrow R$, or $\gamma : R^2 \times R \rightarrow R^2$. A feature vector X is composed as $X = [x^1, x^2, \dots, x^i, \dots, x^I]^T$. The difference between images U_j and U_k with respect to feature vector X and chosen measure d is described by the modulus of vector of distances between corresponding homotypic fragments:

$$|D_{jk}| = |D(U_j, U_k)| = \left| [d_{jk}^1, d_{jk}^2, \dots, d_{jk}^I]^T \right| = \sqrt{\sum_{i=1}^I (d_{jk}^i)^2}, \quad (1)$$

where D_{jk} is an I -dimensional vector of distance between images, $d_{jk}^i = d(u_j^i, u_k^i) = d(x_j^i, x_k^i)$ is a distance between homotypic fragments u_j^i and u_k^i . Let U_l , $l = J + 1$ be an image with unknown attribution. It is necessary to find image U_m (and the author) providing minimum of distance D_{ml} .

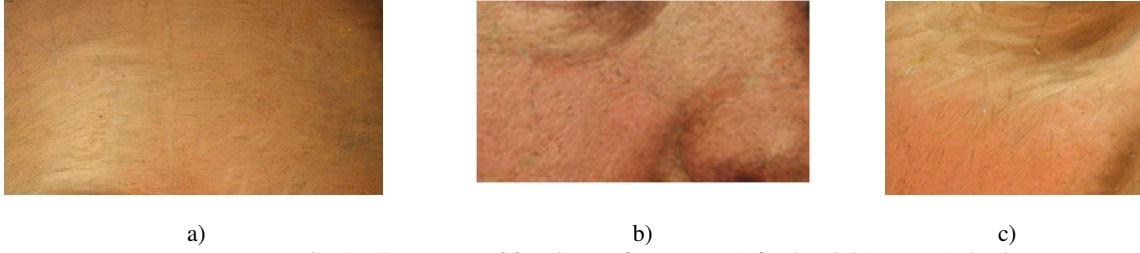


Fig. 1. Three types of face image fragments: a) forehead; b) nose; c) cheek

The workflow includes the following stages: a) finding similar informative fragments in the researched images; b) computing features; c) comparing homotypic fragments; d) formulating attribution decision. The paper deals with stages (b) and (c).

1. Images of paintings

The data used in the research are the image fragments of artworks painted in XVIII – XIX centuries by different authors. The images are fixed by a digital camera. The size of the images is about 4272x2848 pixels. The distortions conditioned by acquisition process are compensated and images are uniformly oriented. The size of informative fragments varies from 990x814 to 1800x1000 pixels at resolution of 200 dots per cm that corresponds to the quality of the data used in the analogous studies. For example, Johnson et al. [5], Polatkan et al. [6], and Li et al. [10] analyzed images obtained at resolution of 196 dots per inch. Some of the paintings have retouched and repainted areas. The features should be extracted only from areas with original brushwork. Thereby, retouched and repainted areas should be excluded during feature extraction. A technique developed in [12] is used for localizing damaged paint layer areas.

2. Image features

In accordance to recommendations of art experts and publications in the subject domain of current research, it is preferable to use image features that do not need segmentation of a single brushstroke. The features will be extracted only from the areas containing maximum information about the artistic manner of the painter. For describing individual manner of artists the following textural features are proposed in [13]: (a) local orientation of grayscale image ridges; (b) simple neighborhood orientation based on the local structure tensor. Histograms of brushstroke ridge orientation and local neighborhood orientation are considered as features of a brushstroke group that describe the individual artistic manner specific to a particular detail of a painting. In this paper, the second of the textural features listed above is considered.

3. Neighborhood orientation

Let the grayscale image relief be a function $f(x, y)$, $f \in C^2(R^2, R)$. A feature describing the local orientation of painting texture is based on the notion of structure tensor, or the second moment matrix at image point x weighted by a window function [15]:

$$\mu_f(x) = \int_{p \in R^2} (Df(p)(Df(p))^T) w(x-p) dp, \quad (2)$$

where $w(x-p)$ is a window Gaussian function [14], $Df = (f_x, f_y)^T$. The angle of local orientation φ is determined as:

$$\varphi = \frac{\pi}{2} + \frac{1}{2} \arctan \frac{2\mu_{f1,1}}{\mu_{f2,0} - \mu_{f0,2}}, \quad (3)$$

where $\mu_{f_i,j}$ are the components of the structure tensor (2):

$$\mu_f = \begin{bmatrix} \mu_{f2,0} & -\mu_{f1,1} \\ -\mu_{f1,1} & \mu_{f0,2} \end{bmatrix}.$$

In [5] the local orientation angle is obtained from the responses of 12 oriented Gabor filters. The second moment matrix based technique used in this work provides an accuracy of angle estimation within 1 degree for a window size of 5 pixels and $\sigma=1$. Histogram of simple neighborhood orientation is obtained directly from an image and does not need brushstroke segmentation. The efficiency of this feature is illustrated below. In fig. 2, the image samples of the same type “nose” taken from artworks of several painters are shown. Corresponding histograms of orientation angle (3) of vectors, orthogonal to the neighborhood orientation vector are shown in fig. 3.

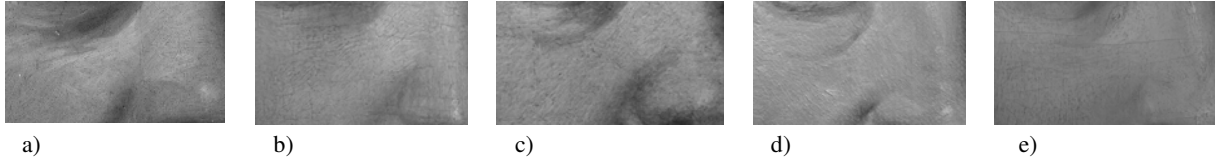


Fig. 2. Homotypic image fragments “nose”

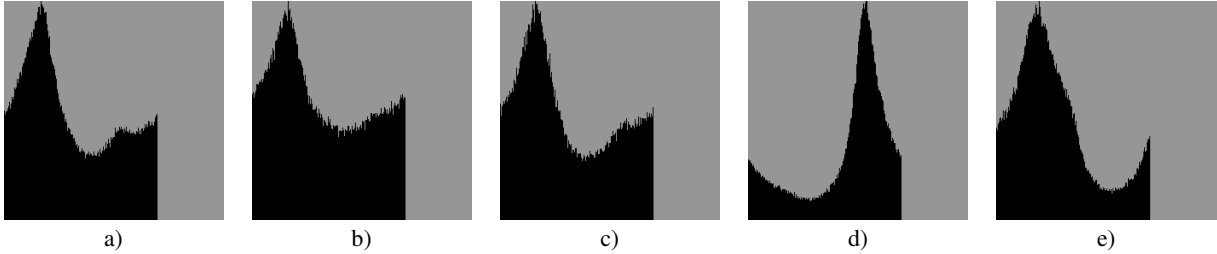


Fig. 3. Histograms of values of angle φ calculated from samples shown in fig. 2

4. Comparing images of paintings

For comparing homotypic informative fragments of artworks the statistical tests [10], cluster analysis, and classification techniques [5, 6] are used. In this paper, the image samples are compared using information-theoretical disparity measure, because this measure fits the features represented by distributions. The measure is constructed on the basis of Kullback-Leibler divergence [16] as follows:

$$d_{jk}^i = d(u_j^i, u_k^i) = \frac{1}{2} \left[\sum_{\varphi \in H} p_{\Phi}(\varphi) \log \frac{p_{\Phi}(\varphi)}{q_{\Phi}(\varphi)} - \sum_{\varphi \in H} q_{\Phi}(\varphi) \log \frac{q_{\Phi}(\varphi)}{p_{\Phi}(\varphi)} \right],$$

where $p_{\Phi}(\varphi)$ and $q_{\Phi}(\varphi)$ are the probabilities of the event when orientation angle values in samples u_j^i and u_k^i are equal to φ ; H is the alphabet of Φ . The measure $d(u_j^i, u_k^i)$ is non-negative and symmetric.

The result of testing features extracted from the homotypic regions (forehead, nose, and cheek) of three portraits by F. Rokotov (see fig. 2, a-c) and two portraits by other artists (see fig. 2 d, e) dated to eighteenth-nineteenth centuries are obtained and presented in fig. 3 and tables 1–3. Here the paintings by F. Rokotov are denoted as R1, R2, and R3. Male and female portraits by other painters are denoted as M and F. The histograms of orientation angle (see fig. 3) extracted from samples shown in fig. 2 demonstrate similarity of homotypic texture regions in portraits by F. Rokotov and dissimilarity of ones being compared to regions of the same type in portraits M and F. The distance between images is calculated using (1) and given in table 4. It can be seen from tables 1-3 that the fragments “nose” and “cheek” in image R3 are closer in terms of measure d_{jk}^i to corresponding fragments in images R1 and R2 than to the fragments in M and F. The fragment “forehead” in R3 is closer to “forehead” in F. Taking into consideration the results given in table 4, one can conclude that the artistic style of images R1, R2, and R3 differs from painting style of two other painters with respect to distance measure (1).

Table 1

Values of disparity measure between fragments "forehead"

Picture	R1	R2	R3	M	F
R1	0	0.009	0.03	0.56	0.049
R2	0.009	0	0.042	0.617	0.08
R3	0.03	0.042	0	0.434	0.013
M	0.56	0.617	0.434	0	0.4
F	0.049	0.56	0.0126	0.4	0

Table 2

Values of disparity measure between fragments "nose"

Picture	R1	R2	R3	M	F
R1	0	0.0067	0.0016	0.449	0.067
R2	0.0067	0	0.0076	0.399	0.038
R3	0.0016	0.0076	0	0.443	0.071
M	0.449	0.399	0.443	0	0.292
F	0.067	0.038	0.071	0.292	0

Table 3

Values of disparity measure between fragments "cheek"

Picture	R1	R2	R3	M	F
R1	0	0.0182	0.0037	0.249	0.087
R2	0.0182	0	0.0166	0.155	0.039
R3	0.0037	0.0166	0	0.233	0.095
M	0.249	0.155	0.233	0	0.173
F	0.087	0.039	0.095	0.173	0

Table 4

Distance between images of paintings

Picture	R1	R2	R3	M	F
R1	0	0.0217	0.0311	0.760	0.1202
R2	0.0217	0	0.0458	0.7506	0.0964
R3	0.0311	0.0458	0	0.6648	0.1193
M	0.760	0.7506	0.6648	0	0.5258
F	0.1202	0.0964	0.1193	0.5258	0

Conclusion

A feature description of images of paintings based on textural characteristics is proposed. Selected image features in the form of orientation angle distributions give quantitative description of a painter artistic style and provide suitable accuracy of features computation. Feature evaluation does not require segmentation of single brushstroke and is not sensitive to image acquisition conditions as opposed to conventional techniques. Results of the computing experiments showed the efficiency of the proposed features for comparing artistic styles. The proposed feature set may be used as a part of technological description of fine art paintings for restoration and attribution. The future research will be aimed at the extension of feature space and creating a procedure for making decisions on similarity of brushwork techniques of the researched paintings based on the extended feature space.

This work is supported by the RFBR grant No 12-07-00668.

References

1. Stork, D. Computer vision and computer graphics analysis of paintings and drawings: an introduction to the literature / D. Stork // LNCS: Springer – 2009. – Vol. 5702. – P. 9–24.
2. Stone, D.M. Computer-assisted Connoisseurship: The Interdisciplinary Science of Computer Vision and Image Analysis in the Study of Art / D.M. Stone, D. Stork // The 3rd Intern. Workshop on Image Processing for Art Investigation IP4AI3. – N.Y., 2010. – P. 9–10.
3. The National gallery glossary [Electronic resource]. – Mode of access : <http://www.nationalgallery.org.uk/paintings/glossary/>. – Date of access : 18.04.2014.
4. Obukhov, G.G. A brief glossary of fine art / G.G. Obukhov. – M., 1959.

5. Image processing for artist identification (Computerized analysis of Vincent van Gogh's painting brushstrokes) / C. R. Johnson [et al.] // *Signal Processing Magazine, IEEE* – 2008. – Vol. 25, No 4. – P. 37–48.
6. Detection of forgery in paintings using supervised learning / S. Polatkan [et al.] // *ICIP2009, IEEE*. – 2009. – P. 2921–2924.
7. Sablatnig R. Structural Analysis of Paintings Based on Brush Strokes / R. Sablatnig, P. Kammerer, E. Zolda // *Proc. of SPIE Scientific Detection of Fakery in Art.* – SPIE. 1998. – V. 3315. – P. 87–98.
8. Lettner, M. Texture Analysis for Stroke Classification in Infrared Reflectograms / M. Lettner, R. Sablatnig // H. Kalviainen et al. Eds., *Proc. Of the SCIA 2005. LNCS.* – Springer-Verlag, Berlin-Heidelberg, 2005. – Vol. 3540. – P. 459–469.
9. Shahram, M. Recovering layers of brush strokes through statistical analysis of color and shape: an application to van Gogh's Self portrait with grey felt hat / M. Shahram, D.G. Stork, D. Donoho // *Computer Image Analysis in the Study of Art. D. Stork, J. Coddington, Eds.: Proc. of the SPIE.* – 2008. – Vol. 6810. – P. 6810D-1–6810D-8.
10. Rhythmic brushstrokes distinguish van Gogh from his contemporaries: findings via automated brushstroke extraction / J. Li [et al.] // *IEEE TPAMI.* – 2012. – Vol. 34, No 6. – P. 1159–1176.
11. Ignatova, N.S. Analysis of oil painting textures / N.S. Ignatova // *Fundamentals of Oil painting Examination.* – M. : Grabar restoration Centre, 1994. – Vol. 1. – P.12–26.
12. Murashov, D.M. Localization of Differences between Multimodal Images on the Basis of an Information-Theoretical Measure / D.M. Murashov // *Pattern Recognition and Image Analysis.* – Springer US, 2014. – Vol. 24, no. 1. – P. 133 – 143.
13. Murashov, D. Composing image feature space for painting attribution tasks / D. Murashov // *Proc. of the 11th Intern. Conference PRIA-11.* – Samara : IPSI RAS, 2013. – Vol. 2. – P. 674–677.
14. Lindeberg, T. *Scale-space Theory in Computer Vision. The Kluwer International Series in Engineering and Computer Science* / T. Lindeberg. – Kluwer Academic Publishers, 1994.
15. Jähne, B. *Digital Image Processing* / B. Jähne. – 6th ed. – Springer, 2005.
16. Escolano, F. *Information Theory in Computer Vision and Pattern Recognition* / F. Escolano, P. Suau, B. Bonev. – London : Springer Verlag, 2009.

DISTANCE-BASED EMBEDDING PROCEDURES FOR PATTERN CLASSIFICATION

N.A. Nechval¹, K.N. Nechval², I. Bausova¹,
G. Berzins¹, J. Krasts¹, V.F. Strelchonok³, M. Moldovan⁴
¹University of Latvia, Riga;
²Transport and Telecommunication Institute, Riga, Latvia;
³Baltic International Academy, Riga, Latvia;
⁴University of New South Wales, Sydney, Australia
e-mail: nechval@junik.lv

In this paper, distance-based embedding procedures for pattern classification (recognition) are proposed which allow one to classify, say, radar clutter into one of several major categories, including bird, weather, and target classes. These procedures do not require the arbitrary selection of priors as in the Bayesian classifier. The decision rule of pattern classification via embedding is based on comparison of distances and is in the form of associating the p -dimensional vector of observations on the object with one of the k specific classes. The results obtained in this paper agree with the simulation results, which confirm the validity of the theoretical predictions of performance of the presented procedures. An application example is given.

Introduction

Classification is often referred to simply as discriminant analysis. In engineering and computer science, classification is usually called pattern recognition. Some writers use the term classification analysis to describe cluster analysis, in which the observations are clustered according to variable values rather than into predefined classes.

In classification, a sampling unit (subject or object) whose group membership is unknown is assigned to a group on the basis of the vector of p measured values, \mathbf{x} , associated with the unit. To classify the unit, we must have available a previously obtained sample of observation vectors from each class. Then one approach is to compare \mathbf{x} with the mean vectors $\bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2, \dots, \bar{\mathbf{x}}_k$ of the k samples and assign the unit to the class whose $\bar{\mathbf{x}}_i$ is closest to \mathbf{x} .

1. Embedding Procedures for Pattern Classification into Two Classes

Classification via Mahalanobis Distance. Let us assume that the two populations have the same covariance matrix ($\Sigma_1 = \Sigma_2$).

If \mathbf{y} has been embedded in the sample from the class C_1 , the Mahalanobis distance between two vectors $\bar{\mathbf{x}}_{\bullet 1}$ and $\bar{\mathbf{x}}_2$ is given by

$$d_{\bullet 12} = (\bar{\mathbf{x}}_{\bullet 1} - \bar{\mathbf{x}}_2)' \mathbf{S}_{\bullet 12}^{-1} (\bar{\mathbf{x}}_{\bullet 1} - \bar{\mathbf{x}}_2). \quad (1)$$

If \mathbf{y} has been embedded in the sample from the class C_2 , the Mahalanobis distance between two vectors $\bar{\mathbf{x}}_1$ and $\bar{\mathbf{x}}_{\bullet 2}$ is given by

$$d_{12\bullet} = (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_{\bullet 2})' \mathbf{S}_{12\bullet}^{-1} (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_{\bullet 2}), \quad (2)$$

where \mathbf{S}_{12} is the pooled within-class covariance matrix, in its bias-corrected form given by

$$\mathbf{S}_{12} = [(n_1 - 1)\mathbf{S}_1 + (n_2 - 1)\mathbf{S}_2] / (n_1 + n_2 - 2), \quad (3)$$

\mathbf{S}_1 and \mathbf{S}_2 are the unbiased estimates of the covariance matrices of the classes C_1 and C_2 , respectively, and there are n_i observations in class C_i ($n_1 + n_2 = n$).

Then the classification rule becomes: Assign \mathbf{y} to C_1 if

$$d_{\bullet 12} > d_{12\bullet}, \quad (4)$$

and assign \mathbf{y} to C_2 if

$$d_{12\bullet} > d_{\bullet 12}. \quad (5)$$

If $(\Sigma_1 = \Sigma_2)$ does not hold, then instead of \mathbf{S}_{12} we use

$$\mathbf{S}_{12}^o = \mathbf{S}_1/n_1 + \mathbf{S}_2/n_2. \quad (6)$$

Remark. If $n_1=n_2=n$, then

$$(n-1)/(n+n-2) = 1/2, \quad (7)$$

so

$$\mathbf{S}_{12}^o = \frac{\mathbf{S}_1}{n_1} + \frac{\mathbf{S}_2}{n_2} = \frac{(n_1-1)\mathbf{S}_1 + (n_2-1)\mathbf{S}_2}{n_1+n_2-2} \left(\frac{1}{n} + \frac{1}{n} \right) = \mathbf{S}_{12} \left(\frac{1}{n} + \frac{1}{n} \right). \quad (8)$$

With equal sample sizes, the large sample procedure is essentially the same as the procedure based on the pooled covariance matrix.

Classification via Generalized Euclidean Distance. Let us assume that the two populations have the same covariance matrix $(\Sigma_1 = \Sigma_2)$.

If \mathbf{x} has been embedded in the sample from C_1 , the generalized Euclidean distance (squared) between two vectors $\bar{\mathbf{x}}_{\cdot 1}$ and $\bar{\mathbf{x}}_2$ is given by

$$\tilde{d}_{\cdot 12} = [(\bar{\mathbf{x}}_{\cdot 1} - \bar{\mathbf{x}}_2)'(\bar{\mathbf{x}}_{\cdot 1} - \bar{\mathbf{x}}_2)] / |\mathbf{S}_{\cdot 12}|. \quad (9)$$

If \mathbf{x} has been embedded in the sample from C_2 , the generalized Euclidean distance between two vectors $\bar{\mathbf{x}}_1$ and $\bar{\mathbf{x}}_{2\cdot}$ is given by

$$\tilde{d}_{12\cdot} = [(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_{2\cdot})'(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_{2\cdot})] / |\mathbf{S}_{12\cdot}|. \quad (10)$$

Then the classification rule becomes: Assign \mathbf{x} to C_1 if

$$\tilde{d}_{\cdot 12} > \tilde{d}_{12\cdot}, \quad (11)$$

and assign \mathbf{x} to C_2 if

$$\tilde{d}_{12\cdot} > \tilde{d}_{\cdot 12}. \quad (12)$$

If $(\Sigma_1 = \Sigma_2)$ does not hold, then instead of \mathbf{S}_{12} we use (6).

Classification via Modified Euclidean Distance. Let us assume that the two populations have the same covariance matrix $(\Sigma_1 = \Sigma_2)$.

If \mathbf{x} has been embedded in the sample from C_1 , the modified Euclidean distance between two vectors $\bar{\mathbf{x}}_{\cdot 1}$ and $\bar{\mathbf{x}}_{12}$ is given by

$$\hat{d}_{\cdot 1} = [(\bar{\mathbf{x}}_{\cdot 1} - \bar{\mathbf{x}}_{12})'(\bar{\mathbf{x}}_{\cdot 1} - \bar{\mathbf{x}}_{12})] / |\mathbf{S}_{\cdot 12}|, \quad (13)$$

where

$$\bar{\mathbf{x}}_{12} = \frac{\sum_{i=1}^2 n_i \bar{\mathbf{x}}_i}{\sum_{i=1}^2 n_i}. \quad (14)$$

If \mathbf{x} has been embedded in the sample from C_2 , the generalized Euclidean distance between two vectors $\bar{\mathbf{x}}_{2\cdot}$ and $\bar{\mathbf{x}}_{12}$ is given by

$$\hat{d}_{2\cdot} = [(\bar{\mathbf{x}}_{2\cdot} - \bar{\mathbf{x}}_{12})'(\bar{\mathbf{x}}_{2\cdot} - \bar{\mathbf{x}}_{12})] / |\mathbf{S}_{12\cdot}|. \quad (15)$$

Then the classification rule becomes: Assign \mathbf{x} to C_1 if

$$\hat{d}_{\cdot 1} + \hat{d}_{2\cdot} > \hat{d}_1 + \hat{d}_2, \quad (16)$$

and assign \mathbf{x} to C_2 if

$$\hat{d}_1 + \hat{d}_2 > \hat{d}_{\cdot 1} + \hat{d}_{2\cdot}. \quad (17)$$

If $(\Sigma_1 = \Sigma_2)$ does not hold, then instead of \mathbf{S}_{12} we use (6).

2. Embedding Procedures for Pattern Classification into Several Classes

Classification via Total Mahalanobis Distance. Let us assume that each of the k populations has the same covariance matrix $(\Sigma_1 = \Sigma_2 = \dots = \Sigma_k)$. The Mahalanobis distance between two vectors $\bar{\mathbf{x}}_i$ and $\bar{\mathbf{x}}_j$, where $i, j \in \{1, 2, \dots, k\}$, $i \neq j$, is given by

$$d_{ij} = (\bar{\mathbf{x}}_i - \bar{\mathbf{x}}_j)' \mathbf{S}_{ij}^{-1} (\bar{\mathbf{x}}_i - \bar{\mathbf{x}}_j). \quad (18)$$

If \mathbf{x} has been embedded in the sample from C_i , the Mahalanobis distance between two vectors $\bar{\mathbf{x}}_{\bullet i}$ and $\bar{\mathbf{x}}_j$ is given by

$$d_{\bullet ij} = (\bar{\mathbf{x}}_{\bullet i} - \bar{\mathbf{x}}_j)' \mathbf{S}_{\bullet ij}^{-1} (\bar{\mathbf{x}}_{\bullet i} - \bar{\mathbf{x}}_j). \quad (19)$$

If \mathbf{x} has been embedded in the sample from C_j , the Mahalanobis distance between two vectors $\bar{\mathbf{x}}_i$ and $\bar{\mathbf{x}}_{\bullet j}$ is given by

$$d_{ij\bullet} = (\bar{\mathbf{x}}_i - \bar{\mathbf{x}}_{\bullet j})' \mathbf{S}_{ij\bullet}^{-1} (\bar{\mathbf{x}}_i - \bar{\mathbf{x}}_{\bullet j}). \quad (20)$$

Let

$$d_r(\mathbf{x}) = \sum_{i=1}^{k-1} \sum_{j=i+1}^k d_{ij}, \quad r \in \{1, 2, \dots, k\} \quad (21)$$

be the total Mahalanobis distance in the case of pattern classification into k classes, where

$$d_{ij} = d_{\bullet ij}, \quad \text{if } i = r, \quad (22)$$

and

$$d_{ij} = d_{ij\bullet}, \quad \text{if } j = r. \quad (23)$$

Then the classification rule becomes: Assign \mathbf{x} to the class C_r , $r \in \{1, 2, \dots, k\}$, for which $d_r(\mathbf{x})$ is largest.

If $(\Sigma_1 = \Sigma_2 = \dots = \Sigma_k)$ does not hold, then instead of \mathbf{S}_{ij} we use

$$\mathbf{S}_{ij}^0 = \mathbf{S}_i/n_i + \mathbf{S}_j/n_j. \quad (24)$$

Classification via Total Generalized Euclidean Distance. Let us assume that each of the k populations has the same covariance matrix $(\Sigma_1 = \Sigma_2 = \dots = \Sigma_k)$. The generalized Euclidean distance between two vectors $\bar{\mathbf{x}}_i$ and $\bar{\mathbf{x}}_j$, where $i, j \in \{1, 2, \dots, k\}$, $i \neq j$, is given by

$$\tilde{d}_{ij} = [(\bar{\mathbf{x}}_i - \bar{\mathbf{x}}_j)'(\bar{\mathbf{x}}_i - \bar{\mathbf{x}}_j)]/|\mathbf{S}_{pl}| = [(\bar{\mathbf{x}}_i - \bar{\mathbf{x}}_j)'(\bar{\mathbf{x}}_i - \bar{\mathbf{x}}_j)]/[(n_1 - 1)\mathbf{S}_1 + \dots + (n_k - 1)\mathbf{S}_k]/[\sum_{i=1}^k n_i - k]. \quad (25)$$

If \mathbf{x} has been embedded in the sample from C_i , the generalized Euclidean distance between two vectors $\bar{\mathbf{x}}_{\bullet i}$ and $\bar{\mathbf{x}}_j$ is given by

$$\tilde{d}_{\bullet ij} = [(\bar{\mathbf{x}}_{\bullet i} - \bar{\mathbf{x}}_j)'(\bar{\mathbf{x}}_{\bullet i} - \bar{\mathbf{x}}_j)]/|\mathbf{S}_{pl(\bullet i)}|. \quad (26)$$

If \mathbf{x} has been embedded in the sample from C_j , the generalized Euclidean distance between two vectors $\bar{\mathbf{x}}_i$ and $\bar{\mathbf{x}}_{\bullet j}$ is given by

$$\tilde{d}_{ij\bullet} = [(\bar{\mathbf{x}}_i - \bar{\mathbf{x}}_{\bullet j})'(\bar{\mathbf{x}}_i - \bar{\mathbf{x}}_{\bullet j})]/|\mathbf{S}_{pl(j\bullet)}|. \quad (27)$$

Let

$$\tilde{d}_r(\mathbf{x}) = \sum_{i=1}^{k-1} \sum_{j=i+1}^k \tilde{d}_{ij}, \quad r \in \{1, 2, \dots, k\}, \quad (28)$$

be the total generalized Euclidean distance in the case of pattern classification into k classes, where

$$\tilde{d}_{ij} = \tilde{d}_{.ij}, \text{ if } i = r, \quad (29)$$

and

$$\tilde{d}_{ij} = \tilde{d}_{ij.}, \text{ if } j = r. \quad (30)$$

Then the classification rule becomes: Assign \mathbf{x} to the class C_r , $r \in \{1, 2, \dots, k\}$, for which $\tilde{d}_r(\mathbf{x})$ is largest.

If $(\Sigma_1 = \Sigma_2 = \dots = \Sigma_k)$ does not hold, then instead of \mathbf{S}_{pl} we use

$$\mathbf{S}^0 = \sum_{i=1}^k \frac{\mathbf{S}_i}{n_i}. \quad (31)$$

Classification via Total Modified Euclidean Distance. Let us assume that each of the k populations has the same covariance matrix $(\Sigma_1 = \Sigma_2 = \dots = \Sigma_k)$. The modified Euclidean distance between two vectors $\bar{\mathbf{x}}_i$ and $\bar{\mathbf{x}}$, $i \in \{1, 2, \dots, k\}$, is given by

$$d_i = [(\bar{\mathbf{x}}_i - \bar{\mathbf{x}})'(\bar{\mathbf{x}}_i - \bar{\mathbf{x}})] / |\mathbf{S}_{pl}|, \quad (32)$$

where

$$\bar{\mathbf{x}} = \sum_{i=1}^k n_i \bar{\mathbf{x}}_i / \sum_{i=1}^k n_i \quad (33)$$

represents the 'overall average'.

If \mathbf{x} has been embedded in the sample from C_i , the modified Euclidean distance between two vectors $\bar{\mathbf{x}}_{.i}$ and $\bar{\mathbf{x}}$ is given by

$$d_{.i} = [(\bar{\mathbf{x}}_{.i} - \bar{\mathbf{x}})'(\bar{\mathbf{x}}_{.i} - \bar{\mathbf{x}})] / |\mathbf{S}_{pl(i)}|, \quad i \in \{1, 2, \dots, k\}. \quad (34)$$

Let

$$d_r(\mathbf{x}) = \sum_{i=1}^k \tilde{d}_i, \quad r \in \{1, 2, \dots, k\}, \quad (35)$$

be the total modified Euclidean distance in the case of pattern classification into k classes, where

$$d_i = d_{.i}, \text{ if } i = r. \quad (36)$$

Then the classification rule becomes: Assign \mathbf{x} to the class C_r , $r \in \{1, 2, \dots, k\}$, for which $d_r(\mathbf{x})$ is largest. If $(\Sigma_1 = \Sigma_2 = \dots = \Sigma_k)$ does not hold, then instead of \mathbf{S}_{pl} we use (31).

3. Application Example

This example is adapted from a study [1] concerned with the detection of hemophilia A carriers. To construct a procedure for detecting potential hemophilia A carriers, blood samples were assayed for two groups of women and measurements on the two variables,

$$x_1 = \log_{10}(\text{AHF activity}) \quad \text{and} \quad x_2 = \log_{10}(\text{AHF-like antigen}) \quad (37)$$

recorded (see table 1 [2]). ("AHF" denotes antihemophilic factor.) The first group of $n_1 = 29$ women were selected from a population of women who did not carry the hemophilia gene. This group was called the *normal* group. The second group of $n_2 = 23$ women was selected from known hemophilia A carriers (daughters of hemophiliacs, mothers with more than one hemophilic son, and mothers with one hemophilic son and other hemophilic relatives). This group was called the *obligatory carriers*. The pairs of observations (x_1, x_2) for the two groups are plotted in fig. 1. Also shown are estimated contours containing 50% and 95% of the probability for bivariate normal distributions centered at $\bar{\mathbf{x}}_1$ and $\bar{\mathbf{x}}_2$,

respectively. Their common covariance matrix was taken as the pooled sample covariance matrix S_{pl} . In this example, bivariate normal distributions seem to fit the data fairly well.

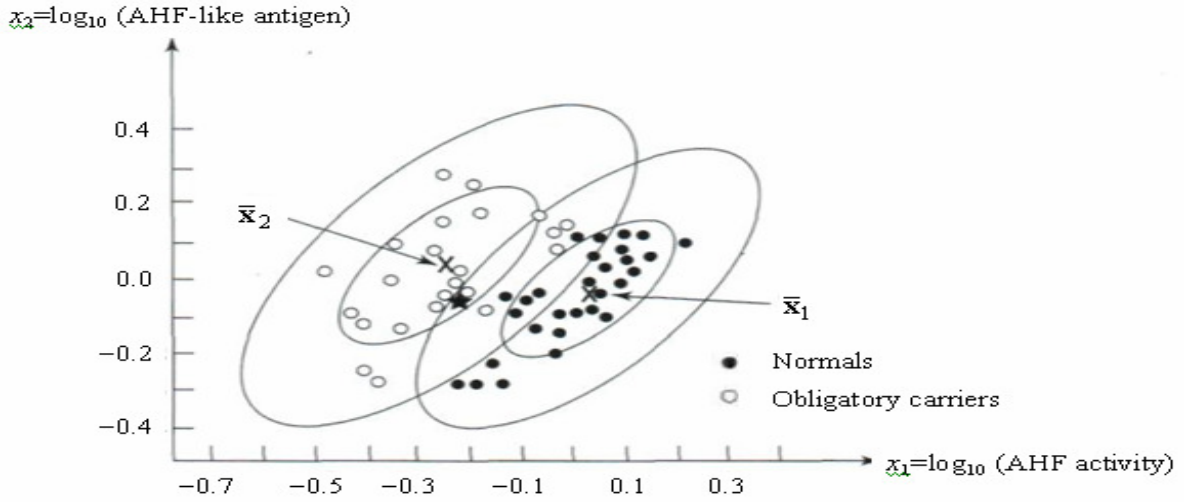


Fig. 1. Image Scatter plots of $[\log_{10}(\text{AHF activity}), \log_{10}(\text{AHF-like antigen})]$ for the normal group and obligatory hemophilia A carriers

The following information is given below:

$$\bar{x}_1 = \begin{bmatrix} 0.00569 \\ -0.04655 \end{bmatrix}, \quad \bar{x}_2 = \begin{bmatrix} -0.24870 \\ 0.01467 \end{bmatrix} \quad \text{and} \quad S_{12}^{-1} = \begin{bmatrix} 132.7928 & -77.5377 \\ -77.5377 & 98.4595 \end{bmatrix}. \quad (38)$$

For instance, measurements of AHF activity and AHF-like antigen on a woman who may be a hemophilia A carrier give $x_1 = -0.210$ and $x_2 = -0.044$ (i.e., $\mathbf{x}' = [x_1, x_2]$). Should this woman be classified as C_1 (normal) or C_2 (obligatory carrier)?

Using Fisher's approach [3], we obtain

$$(\bar{x}_1 - \bar{x}_2)' S_{12}^{-1} \mathbf{x} = -6.96 < (\bar{x}_1 - \bar{x}_2)' S_{12}^{-1} (\bar{x}_1 + \bar{x}_2) / 2 = -4.27. \quad (39)$$

Thus, we classify the woman as C_2 , an obligatory carrier.

Classification via Mahalanobis Distance. It follows from (1) and (2) that

$$d_{\cdot,12} = 9.87, \quad d_{12\cdot} = 11.18. \quad (40)$$

Thus, since

$$d_{12\cdot} > d_{\cdot,12}, \quad (41)$$

we classify the woman as C_2 , an obligatory carrier.

Classification via Generalized Euclidean Distance. It follows from (9) and (10) that

$$\tilde{d}_{\cdot,12} = 425.01, \quad \tilde{d}_{12\cdot} = 486.58. \quad (42)$$

Thus, since

$$\tilde{d}_{12\cdot} > \tilde{d}_{\cdot,12}, \quad (43)$$

we classify the woman as C_2 , an obligatory carrier.

Classification via Modified Euclidean Distance. It follows from (13) and (15) that

$$\hat{d}_1 + \hat{d}_2 = 94.59 + 149.38 = 243.97, \quad \hat{d}_{\cdot,1} + \hat{d}_2 = 77.49 + 150.39 = 227.88. \quad (44)$$

Thus, since

$$\hat{d}_1 + \hat{d}_2 > \hat{d}_{\cdot,1} + \hat{d}_2, \quad (45)$$

we classify the woman as C_2 , an obligatory carrier.

Conclusion

The approaches proposed in this paper represent the improved pattern recognition procedures that allow one to take into account the cases which are not adequate for Fisher's classification rule. Moreover, these approaches allow one to classify sets of multivariate observations, where each of the sets contains more than one observation. For the cases, which are adequate for Fisher's classification rule, the proposed approach gives the results similar to that of Fisher's classification rule.

This research was supported in part by Grant No. 06.1936, Grant No. 07.2036, Grant No. 09.1014, and Grant No. 09.1544 from the Latvian Council of Science and the National Institute of Mathematics and Informatics of Latvia.

References

1. Evaluation of the detection rate of hemophilia carriers / B.N. Bouma [et al.] // *Statistical Methods for Clinical Decision Making*. – 1975. – Vol. 7 (4). – P. 339–350.
2. New approach to pattern recognition via comparison of maximum separations / K.N. Nechval [et al.] // *Computer Modelling and New Technologies*. – 2011. – Vol. 15 (2). – P. 30–40.
3. Fisher, R.A. The use of multiple measurements in taxonomic problems / R.A. Fisher // *Annals of Eugenics*. – 1936. – Vol. 7 (3). – P. 179–188.

STATISTICAL VALIDATION OF SIMULATION MODELS

N.A. Nechval¹, K.N. Nechval², I. Bausova¹, G. Berzins¹,
J. Krasts¹, V.F. Strelchonok³, M. Moldovan⁴

¹ University of Latvia, Riga;

² Transport and Telecommunication Institute, Riga, Latvia;

³ Baltic International Academy, Riga, Latvia;

⁴ University of New South Wales, Sydney, Australia

e-mail: nechval@junik.lv

In this paper, we consider the problem of statistical validation of a multivariate stationary response simulation model of an observed system, which has p response variables. The problem is reduced to testing the equality of the mean vectors for two multivariate normal populations. Without assuming equality of the covariance matrices, it is referred to as the Behrens–Fisher problem. The main purpose of this paper is to bring to the attention of applied researchers a satisfactory test that can be used for testing the equality of two normal mean vectors when the population covariance matrices are unknown and arbitrary. A numerical example is given.

Introduction

Validation is a central aspect to the responsible application of models to scientific and managerial problems. The importance of validation to those who construct and use models is well recognized [1–3]. However, there is little consensus on what is the best way to proceed. This is at least in part due to the variety of models, model applications, and potential tests. The options are manifold, but the guidelines are few.

It is generally preferable to use some form of objective analysis to perform model validation. A common form of objective analysis for validating simulation models is statistical hypothesis testing. Statistical hypothesis testing, as distinguished from graphical or descriptive techniques, offers a framework that is particularly attractive for model validation. A test would compare a sample of observations taken from the target population against a sample of predictions taken from the model. The validity of the model is then assessed by examining the accuracy of model predictions. Such tests have numerous advantages: they provide an objective and quantifiable metric, they are amenable to reduction to a binary outcome, and therefore permit computation of error probability rates, and they accommodate sample-based uncertainty into the test result.

Not surprisingly, a number of statistical tools have been applied to validation problems. For example, Freese [4] introduced an accuracy test based on the standard χ^2 test. Ottosson and Håkanson [5] used R^2 and compared with so-called highest-possible R^2 , which are predictions from common units (parallel time-compatible sets).

The purpose of this paper is to give a methodology based on pivotal quantities for constructing statistical hypothesis tests, which are used for validating a simulation model of a real, observable system.

1. Preliminaries

Detection, decision making, and hypothesis testing are different names for the same procedure. The word detection refers to the effort to decide whether some phenomenon is present or not in a given situation. For example, a radar system attempts to detect whether or not a target is present; a quality control system attempts to detect whether a unit is defective; a medical test detects whether a given disease is present. The meaning has been extended in the communication field to detect which one, among a finite set of mutually exclusive possible transmitted signals, has been transmitted. Decision making is, again, the process of choosing between a number of mutually exclusive alternatives. Hypothesis testing is the same, except the mutually exclusive alternatives are called hypotheses. We usually use the word hypotheses for these alternatives in what follows, since the word seems to conjure up the appropriate intuitive images.

In using statistical hypothesis testing to test the validity of a simulation model under a given experimental frame and for an acceptable range of accuracy consistent with the intended application of the model, we have the following hypotheses:

H_0 : Model is valid for the acceptable range of accuracy under the experimental frame.

H_1 : Model is invalid for the acceptable range of accuracy under the experimental frame.

There are two possibilities for making a wrong decision in statistical hypothesis testing. The first one, type I error, is accepting the alternative hypothesis (H_1) when the null hypothesis (H_0) is actually true, and the second one, type II error, is accepting the null hypothesis when the alternative hypothesis is actually true. In model validation, the first type of wrong decision corresponds to rejecting the validity of the model when it is actually valid, and the second type of wrong decision corresponds to accepting the validity of the model when it is actually invalid. The probability of making the first type of wrong decision is called model builder's risk (α) and the probability of making the second type of wrong decision is called model user's risk (β).

2. Problem Statement

Suppose that we desire to validate a multivariate stationary response simulation model having p normally distributed response variables. Let $\boldsymbol{\mu}_m$ and $\boldsymbol{\mu}_s$ be the population means of the model and system response variable, and let \mathbf{Q}_m and \mathbf{Q}_s be the model and system covariance matrices. We are interested in the testing problem

$$H_0: \boldsymbol{\mu}_m = \boldsymbol{\mu}_s \text{ vs. } H_1: \boldsymbol{\mu}_m \neq \boldsymbol{\mu}_s. \quad (1)$$

Furthermore, let us assume that for the purpose, for which the simulation model is intended, the validity of the model can be determined with respect to its mean response, and the acceptable range of accuracy can be expressed as the difference between the means of the model and the system responses and can be stated as

$$|\boldsymbol{\mu}_m - \boldsymbol{\mu}_s| \leq \boldsymbol{\Delta}, \quad (2)$$

where $\boldsymbol{\Delta}$ is a vector of the largest acceptable differences. Thus, we deal with the multivariate Behrens-Fisher problem, where it is assumed that $\boldsymbol{\mu}_m$ and $\boldsymbol{\mu}_s$ are unknown $p \times 1$ vectors and \mathbf{Q}_m and \mathbf{Q}_s are unknown $p \times p$ positive definite covariance matrices.

3. Statistical Validation with Equal

Covariance Matrices

Let the p -vectors $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_{n_m}$ and $\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_{n_s}$ be independent (potential) random samples from \mathbf{X} (model) and \mathbf{Y} (system), respectively, where $\mathbf{X} \sim N_p(\boldsymbol{\mu}_m, \mathbf{Q}_m)$ and $\mathbf{Y} \sim N_p(\boldsymbol{\mu}_s, \mathbf{Q}_s)$. The problem is to compare the mean vectors $\boldsymbol{\mu}_m$ and $\boldsymbol{\mu}_s$.

Here we consider the case when the covariance matrices of the two normal distributions are equal (say $\mathbf{Q}_m = \mathbf{Q}_s = \mathbf{Q}$). In this case, Hotelling's T^2 statistic is used for testing the hypothesis $H_0: \boldsymbol{\mu}_m = \boldsymbol{\mu}_s$ vs. $H_1: \boldsymbol{\mu}_m \neq \boldsymbol{\mu}_s$.

Define

$$\bar{\mathbf{X}} = \frac{1}{n_m} \sum_{i=1}^{n_m} \mathbf{X}_i, \quad \bar{\mathbf{Y}} = \frac{1}{n_s} \sum_{j=1}^{n_s} \mathbf{Y}_j, \quad \mathbf{S}_m = \frac{1}{n_m - 1} \sum_{i=1}^{n_m} (\mathbf{X}_i - \bar{\mathbf{X}})(\mathbf{X}_i - \bar{\mathbf{X}})', \quad \mathbf{S}_s = \frac{1}{n_s - 1} \sum_{j=1}^{n_s} (\mathbf{Y}_j - \bar{\mathbf{Y}})(\mathbf{Y}_j - \bar{\mathbf{Y}})'. \quad (3)$$

Then

$$\bar{\mathbf{X}} \sim N_p(\boldsymbol{\mu}_m, (\mathbf{Q}_m / n_m)), \quad \bar{\mathbf{Y}} \sim N_p(\boldsymbol{\mu}_s, (1/n_s)\mathbf{Q}_s), \quad (n_m - 1)\mathbf{S}_m \sim W_p(n_m - 1, \mathbf{Q}_m), \quad (n_s - 1)\mathbf{S}_s \sim W_p(n_s - 1, \mathbf{Q}_s), \quad (4)$$

and these are mutually independent, where $W_p(n, \mathbf{Q})$ denotes the Wishart distribution with n degrees of freedom and corresponding covariance matrix \mathbf{Q} . If $\mathbf{Q}_m = \mathbf{Q}_s = \mathbf{Q}$, then

$$\bar{\mathbf{X}} - \bar{\mathbf{Y}} \sim N_p(\boldsymbol{\mu}_m - \boldsymbol{\mu}_s, (1/n_m + 1/n_s)\mathbf{Q}), \quad (n_m - 1)\mathbf{S}_m + (n_s - 1)\mathbf{S}_s = (n_m + n_s - 2)\mathbf{S} \sim W_p(n_m + n_s - 2, \mathbf{Q}), \quad (5)$$

where “ \sim ” denotes “is distributed as”. Thus the test statistic is

$$T^2 = (\bar{\mathbf{X}} - \bar{\mathbf{Y}})' [(1/n_m + 1/n_s)\mathbf{S}]^{-1} (\bar{\mathbf{X}} - \bar{\mathbf{Y}}) = \frac{n_m n_s}{n_m + n_s} (\bar{\mathbf{X}} - \bar{\mathbf{Y}})' \mathbf{S}^{-1} (\bar{\mathbf{X}} - \bar{\mathbf{Y}}). \quad (6)$$

The T^2 statistic has the central T^2 distribution when $\boldsymbol{\mu}_m = \boldsymbol{\mu}_s$ is true. For T^2 to have an F distribution, the expression for T^2 must be weighted by the factor $(n_m + n_s - p - 1)/[p(n_m + n_s - 2)]$ so that

$$F = \frac{n_m + n_s - p - 1}{p(n_m + n_s - 2)} T^2 \sim F_{p, n_m + n_s - p - 1}, \quad (7)$$

where $F_{p, n_m + n_s - p - 1}$ is the F distribution with degrees of freedom p and $n_m + n_s - p - 1$. If

$$F \leq F_{p, n_m + n_s - p - 1; 1 - \alpha} \text{ or } T^2 \leq \frac{p(n_m + n_s - 2)}{n_m + n_s - p - 1} F_{p, n_m + n_s - p - 1; 1 - \alpha}, \quad (8)$$

where

$$\Pr\{F \leq F_{p, n_m + n_s - p - 1; 1 - \alpha}\} = 1 - \alpha, \quad (9)$$

the null hypothesis H_0 (1) is accepted and rejected otherwise.

A $100(1 - \alpha)\%$ confidence region for $\boldsymbol{\mu}_m - \boldsymbol{\mu}_s$ is given by

$$C_{\boldsymbol{\mu}_m - \boldsymbol{\mu}_s}, \quad (10)$$

where

$$V = \frac{n_m + n_s - p - 1}{p(n_m + n_s - 2)} \frac{n_m n_s}{n_m + n_s} [(\bar{\mathbf{X}} - \bar{\mathbf{Y}}) - (\boldsymbol{\mu}_m - \boldsymbol{\mu}_s)]' \quad (11)$$

represents a pivotal quantity whose distribution does not depend on unknown parameters.

Thus, the decision rule for testing the validity of the model with specified model user's risk (α) is the following: Accept the validity of the model for the acceptable range of accuracy (2) under the given experimental frame if (8) takes place and

$$|\boldsymbol{\mu}_m - \boldsymbol{\mu}_s| \leq \Delta \text{ for all } (\boldsymbol{\mu}_m - \boldsymbol{\mu}_s) \in C_{\boldsymbol{\mu}_m - \boldsymbol{\mu}_s}. \quad (12)$$

Testing the equality of two covariance matrices. For testing the hypothesis $H_0: \mathbf{Q}_m = \mathbf{Q}_s$, for two multivariate normal populations, the statistic

$$M = (n_m + n_s - 2) \ln |\mathbf{S}| \quad (13)$$

is used. χ^2 approximation for the probability distribution of M is given by Box [6] as

$$M(1 - A) \sim \chi_f^2, \quad (14)$$

where

$$f = \frac{p(p+1)}{2} \quad (15)$$

is the number of degrees of freedom,

$$A = \frac{2p^2 + 3p - 1}{6(p+1)} \left(\frac{1}{n_m - 1} + \frac{1}{n_s - 1} - \frac{1}{n_m + n_s - 2} \right). \quad (16)$$

The hypothesis H_0 is rejected if

$$M(1 - A) > \chi_{f; 1 - \alpha}^2, \quad (17)$$

where α is the significance level.

4. Statistical Validation with Unequal

Covariance Matrices

Let the p -vectors $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_{n_m}$, and $\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_{n_s}$ be independent (potential) random samples from \mathbf{X} (model) and \mathbf{Y} (system), respectively, where $\mathbf{X} \sim N_p(\boldsymbol{\mu}_m, \mathbf{Q}_m)$ and $\mathbf{Y} \sim N_p(\boldsymbol{\mu}_s, \mathbf{Q}_s)$. The problem is to compare the mean vectors $\boldsymbol{\mu}_m$ and $\boldsymbol{\mu}_s$.

In this section, we consider the case when the covariance matrices of the two normal distributions are unequal (i.e., $\mathbf{Q}_m \neq \mathbf{Q}_s$). In this case, for testing the hypothesis $H_0: \boldsymbol{\mu}_m = \boldsymbol{\mu}_s$ vs. $H_1: \boldsymbol{\mu}_m \neq \boldsymbol{\mu}_s$, a very natural statistic

$$T_{\bullet}^2 = (\bar{\mathbf{X}} - \bar{\mathbf{Y}})' [\mathbf{S}_m / n_m + \mathbf{S}_s / n_s]^{-1} (\bar{\mathbf{X}} - \bar{\mathbf{Y}}) \quad (18)$$

can be used. This test statistic has approximately Hotelling's distribution, as it is shown in Krishnamoorthy and Yu [7], i.e.,

$$T_{\bullet}^2 \sim \frac{vp}{v-p+1} F_{p, v-p+1} \text{ approximately,} \quad (19)$$

where

$$v = p(p+1) \quad (20)$$

Krishnamoorthy and Yu [7] showed that v is bound in the same way as in the one-dimensional case,

$$\min(n_m - 1, n_s - 1) \leq v \leq n_m + n_s - 2, \quad (21)$$

v being close to the upper bound tells us that the two variance matrices are (almost) equal. The closer v is to the lower bound, the bigger the discrepancy is between them. The lower bound is attained only if one of $\mathbf{S}_m, \mathbf{S}_s$ is a zero matrix.

A $100(1-\alpha)\%$ confidence region for $\boldsymbol{\mu}_m - \boldsymbol{\mu}_s$ is given by

$$C_{\boldsymbol{\mu}_m - \boldsymbol{\mu}_s} = \left\{ \begin{array}{l} \boldsymbol{\mu}_m - \boldsymbol{\mu}_s : [(\bar{\mathbf{x}} - \bar{\mathbf{y}}) - (\boldsymbol{\mu}_m - \boldsymbol{\mu}_s)] \left(\frac{\mathbf{S}_m}{n_m} + \frac{\mathbf{S}_s}{n_s} \right)^{-1} \\ \times [(\bar{\mathbf{x}} - \bar{\mathbf{y}}) - (\boldsymbol{\mu}_m - \boldsymbol{\mu}_s)] \\ \leq \frac{pV}{v-p+1} F_{p, v-p+1; 1-\alpha} \end{array} \right\}. \quad (22)$$

Thus, in this case, the decision rule for testing the validity of the model with specified model builder's risk is the following: Accept the validity of the model for the acceptable range of accuracy (2) under the given experimental frame if

$$T_{\bullet}^2 \leq \frac{pV}{v-p+1} F_{p, v-p+1; 1-\alpha} \quad (23)$$

and (22) takes place.

5. Numerical Example

Based on the experimental data from the simulation model ($n_m=28$) and the observed system ($n_s=28$), the summary statistics for the simulation model (subscript m) and the observed system (subscript s) follow

$$\bar{\mathbf{X}} = \begin{bmatrix} 29.143 \\ 48.643 \\ 35.571 \\ 86.500 \end{bmatrix}, \quad \bar{\mathbf{Y}} = \begin{bmatrix} 28.964 \\ 45.179 \\ 34.679 \\ 81.964 \end{bmatrix} \quad \text{and} \quad \bar{\mathbf{X}} - \bar{\mathbf{Y}} = \begin{bmatrix} 0.179 \\ 3.464 \\ 0.893 \\ 4.536 \end{bmatrix}. \quad (24)$$

The sample covariance matrices defined in (3) are:

$$\mathbf{S}_m = \begin{bmatrix} 22.942 & 30.942 & 4.434 & 21.815 \\ 30.942 & 78.608 & 14.582 & 56.704 \\ 4.434 & 14.582 & 17.513 & 30.519 \\ 21.815 & 56.704 & 30.519 & 91.074 \end{bmatrix} \text{ and } \mathbf{S}_s = \begin{bmatrix} 24.036 & 18.747 & 15.062 & 31.517 \\ 18.747 & 42.374 & 11.726 & 38.451 \\ 15.062 & 11.726 & 20.522 & 31.951 \\ 31.517 & 38.451 & 31.951 & 132.258 \end{bmatrix}. \quad (25)$$

Consider testing

$$H_0: \boldsymbol{\mu}_m - \boldsymbol{\mu}_s = 0 \text{ vs. } H_1: \boldsymbol{\mu}_m - \boldsymbol{\mu}_s \neq 0. \quad (26)$$

Statistical validation when $\mathbf{Q}_m = \mathbf{Q}_s$. The pooled sample covariance matrix is given by

$$\mathbf{S} = \begin{bmatrix} 23.489 & 24.845 & 9.748 & 26.666 \\ 24.845 & 60.491 & 13.154 & 47.577 \\ 9.748 & 13.154 & 19.018 & 31.235 \\ 26.666 & 47.577 & 31.235 & 111.666 \end{bmatrix} \quad (27)$$

and the Hotelling T^2 statistic defined in (6) is computed as

$$\text{Taking } \alpha = 0.025, \text{ we have that } T^2 = (\bar{\mathbf{X}} - \bar{\mathbf{Y}})'[(1/n_m + 1/n_s)\mathbf{S}]^{-1}(\bar{\mathbf{X}} - \bar{\mathbf{Y}}) = 5.646. \quad (28)$$

$$\frac{p(n_m + n_s - 2)}{n_m + n_s - p - 1} F_{p, n_m + n_s - p - 1; 1 - \alpha} = 10.814. \quad (29)$$

Because $T^2 = 5.646 < 12.93$, we do not reject H_0 . Thus, the data do not provide sufficient evidence to indicate that the mean vectors are significantly different.

Testing the equality of two covariance matrices. It follows from (13) and (16) that

$$M(1 - A) = 19.25377 < \chi_{f; 1 - \alpha}^2 = 20.48, \quad (30)$$

where $M=20.91959$, $A=0.07963$, $f=10$, $\alpha=0.025$. Thus, the data do not provide sufficient evidence to indicate that the covariance matrices are significantly different.

Statistical validation when $\mathbf{Q}_m \neq \mathbf{Q}_s$. For this example, the values of T^2 and T_*^2 must be the same because the sample sizes are equal. Thus,

$$T_*^2 = (\bar{\mathbf{X}} - \bar{\mathbf{Y}})'[\mathbf{S}_m/n_m + \mathbf{S}_s/n_s]^{-1}(\bar{\mathbf{X}} - \bar{\mathbf{Y}}) = 5.646. \quad (31)$$

The approximate degree of freedom

$$v = p(p + 1) \quad (32)$$

and the critical value

$$\frac{pv}{v - p + 1} F_{p, v - p + 1; 1 - \alpha} = 10.868. \quad (33)$$

Because $T_*^2 = 5.646 < 12.988$, we do not reject H_0 .

Thus, both tests produced similar results and yielded the same conclusions.

Conclusion

In this paper, the definition of a confidence region is generalized so that problems such as constructing exact confidence regions for the difference in two multivariate normal means can be tackled without the assumption of equal covariance matrices. This allows one to provide satisfactory solutions in a variety of problems, not just the ones reported here.

This research was supported in part by Grant No. 09.1544 from the Latvian Council of Science.

References

1. Balci, O. A bibliography on the credibility, assessment and validation of simulation and mathematical models / O. Balci, R.G. Sargent // *Simuletter*. – 1984. – Vol. 15. – P. 15–27.
2. Nechval, K.N. Technique for identifying an observable process with one of several simulation models / K.N. Nechval, N.A. Nechval, E.K. Vasermanis // *Proceedings of the Summer Computer Simulation Conference (SCSC'03)*, Montreal, Canada, July 20–24, 2003. – Montreal, 2003. – P. 70–75.
3. Vasermanis, E.K. Statistical validation of simulation models of observable systems / E.K. Vasermanis, K.N. Nechval, N.A. Nechval // *Kybernetes (The International Journal of Systems & Cybernetics)*. – 2003. – Vol. 32. – P. 858–869.
4. Freese, F. Testing accuracy / F. Freese // *Forest Sci.* – 1960. – Vol. 6. – P. 139–145.
5. Ottosson, F. Presentation and analysis of a model simulating the pH response of lake liming / F. Ottosson, L. Håkanson // *Ecol. Modelling*. – 1997. – Vol. 104. – P. 89–111.
6. Box, G.E.P. A general distribution theory for a class of likelihood criteria / G.E.P. Box // *Biometrika*. – 1994. – Vol. 36. – P. 317–346.
7. Krishnamoorthy, K. Modified Nel and Van der Merwe test for the multivariate Behrens-Fisher problem / K. Krishnamoorthy, J. Yu // *Statistics & Probability Letters*. – 2004. – Vol. 66. – P. 161–169.

SYSTEM FOR AUTOMATIC PARALLELISATION OF PROCESSING LARGE VOLUMES OF IMAGES USING MULTI-USER RESOURCES

A. Nedzved¹, A. Belotserkovsky¹, A. Khmelkov²

¹United Institute of Informatics Problems of the NAS of Belarus, Minsk;

²Belarusian State University, Minsk

e-mail: nedzveda@tut.by

In this paper, we considered a problem of specific software development for parallel processing of large volume of image. It is based on concept of generating a list of tasks and distributed at different computational resources. The effectiveness of such a system is gained due to interaction between database, command interpreter for image analysis and monolithically compiled programs.

Introduction

It is quite complicated to replicate in computerized applications tools for automated image analysis if we need a full universal support of workplace in several ways [1]. Thus, we come to the necessity of using the already popular notion of Software Flexibility. It mustn't be confused with a flexible development methodology – “Agile software development”, – the concept, which is quite popular as a new feature of software applications. The generally accepted definition of software flexibility has been given by IEEE [2] (Flexibility: “The ease with which a system or component can be modified for use in applications or environments other than those for which it was specifically designed”). It can be challenged on a number of positions. In addition there is already a number of metrics at the moment, characterizing the degree of flexibility of the developed environment [3] but sometimes it contradict with each other. This issue is actually not the subject of this paper, since authors are pursuing the global goal to offer the architecture of applied task-driven system that would later allow:

- aiding image analysis, by making the results quantitative, reproducible, and objective;
- selecting relevant subsets of features in diagnostic tasks, by tracing the set of features most commonly used by experts;
- discovering new knowledge, by allowing the extraction of novel features and the execution of statistical tests for evaluating their significance;
- supporting decision making since it can be considered as a first step in the development of a knowledge base for automated analysis and diagnosis.

At this stage, we propose conceptual schemes for process management of image processing system (as an application domain), and generating a user interface, with the intellectual, social (collaborative, recommending) component according to specific needs of certain users and developers. It allows modifications and building-up the functionality due to the wide use of workflows (processing scripts), which obviously will contribute to solve a problems described above.

1. Properties of Image Analysis Software

As of applied needs there are plenty of modern software for image analysis already developed at the moment. In sense of way they were built we can distinguish six types of applications:

- application programming interface (**API**) – a set of functions or routines that are designed to solve a specific task or are allowed to interact with a specific software component (for example VIGRA, IM);
- **script application** – number of instructions (a program) written using script paradigm for special run-time environment that can interpret (rather than compile) and automate the execution of tasks which could alternatively be executed one-by-one (for example luaCV scripts, TCL-ITK);
- **plug-in** – a software component that adds a specific feature to an existing software application (for example ImageJ);
- software **component** – is a software package, a web service, a web resource, or a module that encapsulates a set of related functions (or data) (for example toolboxes for Matlab, Matcad);
- **event component** are handled synchronously with the program flow, that is, the program has one or more dedicated places where events are handled, frequently an event loop (for example IMAQ);
- **source code** is any collection of computer instructions (possibly with comments) written using some human-readable computer language (OpenCV, ITK).

Though these types are mixed for different software packages (for instance, libraries IM, OpenCV may be classified as API and source code also), each type of software has certain differences and properties. They either offer a high degree of *adaptability* with little or no *assurance over consistency* or severely restrict change to achieve consistency (fig. 1). Source code, depicted in the lower right corner, offers complete adaptability because it permits third parties to change the applications source code, its principal behavioral specification. In theory, any change is possible. In practice, the burdens of understanding a source code good enough to make changes coupled with verifying the correctness of those changes poses significant hurdles, especially for large intricate software systems. While any change is possible, few if any assurances are provided over those changes. In response, various strategies have been devised that help raise confidence in source code changes, and hence an applications consistency. These include confining change to certain source files and verifying application invariants using embedded assertion statements or test scripts. The consistency gained using these measures cuts off certain avenues of adaptability.

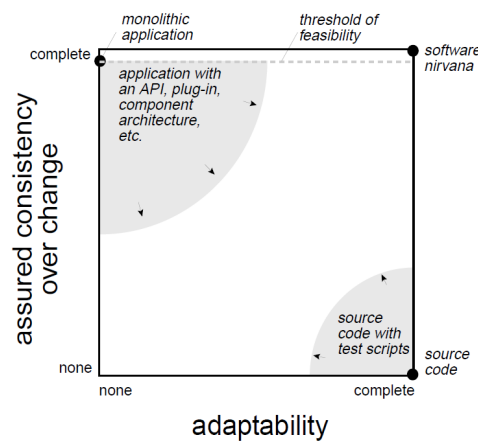


Fig. 1. Dependency between adaptability and consistency offered by current realization techniques for image analysis software

Monolithic applications, which lack such mechanisms, represent the opposite extreme. By prohibiting third party change, application developers can guarantee a high degree of consistency. Incorporating techniques of developing into these applications, by utilizing APIs or component architectures for example, increases adaptability by permitting third parties to change certain aspects of the application, but simultaneously introduces the possibility that those changes could violate consistency.

2. Tools for Image Analysis

At first, data are processed using temporary file, which allows us to analyze records (images). Then, a transfer by calling a run-time library is performed. The system was designed as the main module. It provides interaction of complex components as it is built on the interpreter of Lua language

Lua is a register-based virtual machine. Traditionally, most virtual machines intended for actual execution are stack-based ones. Algorithms, designed with Lua, are optimized by tables and used as arrays: unlike other scripting languages, Lua does not offer an array type. Instead of it Lua programmers use regular tables with integer indices to implement arrays. Lua uses a new algorithm that detects whether tables are being used as arrays. These algorithms automatically store the values associated to numeric indices in an actual array, instead of adding them to the hash table (fig. 2). This mechanism uses an array-based stack to store activating records. There is a novel approach to function closures that keeps local variables in the array-based stack and pushes them to the heap. Every package included in Lua is represented in such table-array.

We set two additional columns in this table to define vector of probability coefficient of image characteristics and vector of coefficient of the target. In this case the target is considered as characteristics of the image processed by current function. These vectors allow to select image processing function on a base of characteristics analysis.

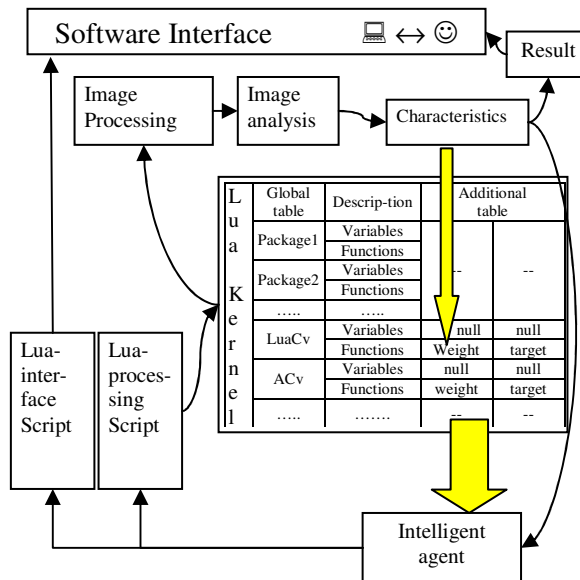


Fig. 2. Script generation cycle for image processing and analysis

All vectors are collected by comparing with image characteristics by means of Lua-table. This collection is sending then to *Intelligent Agent* to generate a new script. Script processes the image and changes characteristics. These characteristics are pushed to Lua-table then to generate a new collection for Intelligent Agents and to represent in User Interface together with processed image. In result software modifies itself and changes processing properties and interface. Image structure is determined by a module of graphical interface into OpenCV library, which is responsible for visualization and representation of images. Headers of image structures are global variables-pointer of interpreter LUA and have special type – user data. User data correspond to pointer in computer address space. This module also includes image read/write function, simple functions of image processing and interactive contouring. All interactive functions return values in the event block, which changes global variables of interpreter. For tasks of monitoring space-occupying lesion a simultaneous usage of several modules is required. In this case, an interaction has been performing by using global variables of the Lua interpreter and properties of user data type.

3. Multitasking system for big data image processing tools

A principle scheme to process and analyze large image volumes based on task tables is proposed was proposed were each of computer systems has its own priority for any task . At every step, functions are indicated by priorities. For instance, for image enhancement the higher priority is defined to noise removing, next priority level includes contrast enhancement and correction of the borders. Priorities determine the order of functions application and necessity of re-analysis using Neural Network. The variable of priority determines location of function in the generated script. In consequence such software has intelligent self-programming possibilities. the concept is based on combination of dynamic libraries, databases and the interpreter features with a set of image processing functions. As a result, the software can be divided into two parts: the first one is focused on an independent application design, and the second one is related to applied users. The interpreter has an opportunity to use additional functions of dynamic libraries. It allows to change software properties and avoid the compilation stage. On the other hand, users can change a task in a database, which allows to parallelize and control its execution process.

It is known that idea of parallel computing is based on the fact that the problems can often be divided into a set of smaller problems, which can be solved simultaneously. Parallel computing is presented in several forms: the bit level parallelism, instruction level parallelism, data parallelism, and task parallelism. In our case, data parallelism will be used. The main idea of the approach based on the data parallelism is that of one operation being performed on all the data array elements. Various fragments of such an array are processed using different processors being nearby or distributed in a computer network. The data distribution between the processors is carried out by the software.

In common, system is constructed from four basic software components: file server for images set, data base for tasks management, manager for generating sets of tasks and manager for task execution (fig. 3).

File server stores all images (original one and processed). All managers are realized as client application. The manager of tasks generating spends estimation of images from file server. Evaluation of image marks is performed based on a specialized tool as described above. This manager creates a list of tasks by linkages image estimation marks with commands of image analysis and processing solutions and user decision. This list is stored in the database. Simple manager of tasks execution takes first command with special status form database and push it to execution. It changes status of command for every stage of processing. In result sets of images are processed by different computing resources at the same time. It decreases a computational cost of image analysis. Parallel computing systems are physical computing and the software systems performing, in one way or another, parallel data processing on many computational nodes. On each processor of the multiprocessor system, a single-threaded process is implemented which communicates with the processes running on other processors using the messages. One sequential function that processes one image (or group of the coupled images) is developed, and the system provides application of this function to large image volume in a parallel mode. It allows to focus on developing the image processing algorithms while avoiding the distraction related to parallel processing implementation.

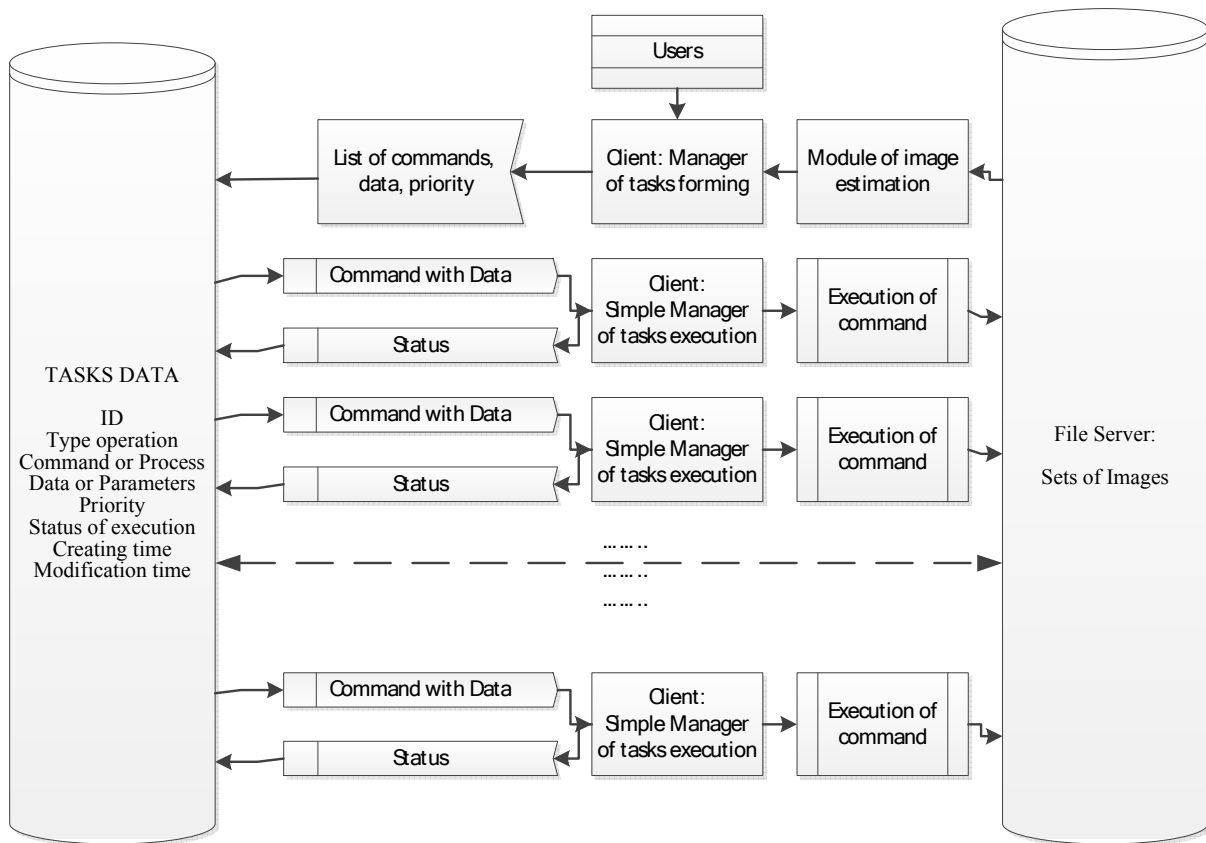


Fig. 3. Information flows of multitasking system for big data image analysis

Conclusion

The technology of software development based on the automatic generation of the scripts including the set of simple image processing functions for a variety of problems of large image sequences analysis. This software is based on the open architecture principles and allows to change the design on the user workstation without compilation stage on the real-time basis. At the same time, the program execution time remains the same as the one in the compiled version. The software dynamism is significantly improved as using the interpreter components, the software can be easily modified and adapted to solving new problems. In addition, the overall concept of such software development will definitely show its worth in complex software systems designing in the nearest future.

References

1. Self-Developing Software for Image Processing Tasks / A. Nedzvedz [et al.] // Proc. on Open German-Russian Workshop "Pattern Recognition and Image Understanding". – Nizhny Novgorod, 2011. – P. 215–218.

2. IEEE. Standard Glossary of Software Engineering Terminology 610.12-1990. – Los Alamitos : IEEE Press, 1999. – Vol. 1.
3. Eden, A.H. Measuring Software Flexibility / A.H. Eden, T. Mens // IEE Software. – London, UK : The Institution of Engineering and Technology, 2006. – Vol. 153, no. 3. – P. 113–126.
4. Oreizy, P. Open Architecture Software: A Flexible Approach to Decentralized Software Evolution / P. Oreizy. – Irvine : University of California, 2000. – 142 p.
5. Gurevich, I.B. Fundamental Concepts and Elements of Image Analysis Ontology / I.B. Gurevich, O. Salvetti, Y. Trusova // Pattern Recognition and Image Analysis: Advances in Mathematical Theory and Applications. – Pleiades Publishing, Ltd., 2009. – Vol. 19, no. 4. – P. 603–611.
6. Bradski, G. Learning OpenCV: Computer Vision with the OpenCV Library / G. Bradski, A. Kaehler. – O'Reilly, 2008.
7. Ierusalimschy, R. Programming in Lua / R. Ierusalimschy. – 2nd ed. – Lua.org, 2006.

DETECTION PATTERNS OF LIVING CELLS FROM ITS AGGREGATION ON THE DIGITAL IMAGE

O. Nedzvedz¹, S. Ablameyko²

¹United Institute of Informatics Problems of the NAS of Belarus, Minsk;

²Belarusian State University, Minsk

e-mail: nedzveda@tut.by

One of the important problem of modern direction of cytological image analysis is cells segmentation. There are many algorithms in this field. But image properties and methods of cell investigation are changing every day. Today most perspective direction of cytological image analysis is living cells investigation. Such images lead to many troubles for cells analysis. In this paper we are proposed solution of one such problems: pattern extraction of cells from its aggregation.

Introduction

Investigation of the features of processes in cell cultures is essential for assessment and prediction of activity of a number of pathological processes. The ability to control these processes is connected to the development of appropriate technical equipment of monitoring of dynamical changes of cells and their biochemical properties in different phases of culture growth. At the same time, the quantitative interpretation of the dynamic changes of cell populations and clusters of cells may allow diagnosis of processes in different tissues of the body at an earlier stage of their development.

Quality monitoring of cell activity is performed in special boxes that significantly degrade the image. On the other hand cells are grouped together and stick together making it difficult to study them. This conditions leads to the fact that for the characterization of the cells during the process of their work requires the development of additional algorithms for their separation and image enhancement.

1. Properties of microscopic images acquisition

Microscope is an instrument used to produce magnified images of objects that are too small to be seen with the naked or unaided eye. Types of microscopes designed to study of microorganisms include the light microscope, dark field microscope, phase contrast microscope, confocal microscope, interference microscope, fluorescence microscope, electron microscope, and atomic force microscope.

Optical microscope (or compound microscope) consists of two lenses: an objective and an eyepiece (ocular). Light microscope uses a beam of light to create an enlarged image of the specimen.

Ray tracings in the optical microscope are the respective objective and eyepiece focal points. The object deform direction of rays. In result maximum intensity of light is concentrate near body center and borders of cells image are usually dark. The real image of the cell formed by the objective on camera matrix and every time support by this properties of cells borders (fig.1). Such image correspond to distortion of rays by cell and don't support information about density of matter in the cell.

The ability of an optical microscope to view an object depends on the size of the object relative to the wavelength of the light used to observe it. Sizes of small details that can be discerned through a microscope depend on the restrictions due to light diffraction. Diffraction is a deviation of light from the rectilinear propagation as it passes around the edge of an object that is physically the approximate size of, or even smaller than that light's wavelength.

Therefore the most important feature for a cell in the image is the shadow. Despite recent advances in modeling the Shape-from-Shading problem and its numerical solution, practical applications have been limited. This is primarily due to the lack of perspective Shape-from-Shading models without the assumption of a light source at the camera center and the non-metric spatial localization of the reconstructed shape. We try a modified formulation of the problem that was described in [1]. That allows the reconstruction of surfaces lit by a near point light source away opposite the camera center. The knowledge of the light source position can enable the recovery of depth information in a metric space by triangulating specular highlights. But such methods lead to many problems of border condition on the image and time required.

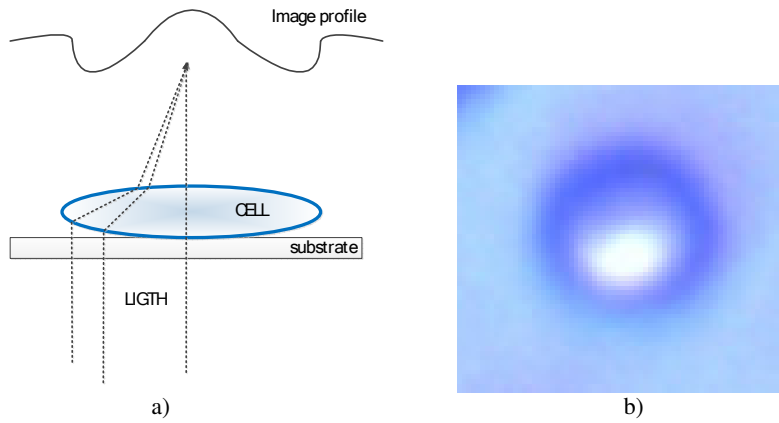


Fig. 1. Properties of images acquisition by microscope: a) light ways for image forming; b) sample of microscopic image of cell

For solution of this problem we are developed simple algorithm for volume reconstruction of cells on the image.

2. Volume reconstruction of cells from image

Usually acquisition of images by optical microscope reflects all distortion optical way of equipment. Therefore light maximum on the image of cell is usually shifted. Cells are aggregate together and change image. In this case classical methods of geometrical reconstruction don't work. But we can use of properties of cell border. Every time the border region is corresponding to local minimum on profile line (fig. 2)

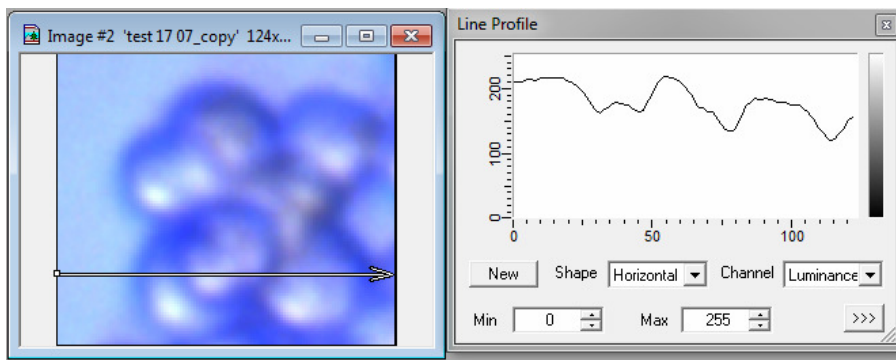


Fig. 2. Image of cells aggregation with profile line of intensity

For easy image understanding the algorithm invert intensity of image and for cells border intensity increase or decrease.

The algorithm is divided on tree stages. Those stages are corresponded to procedures of raster scans, but direction of scanning is changed for every iteration. The first iteration is completed from left to right and top to bottom direction. Every step of this iteration changed of intensity by low:

$$I_{x,y} = \max(I_{x,y}, I_{x-1,y}, I_{x,y-1}, I_{x-1,y-1}),$$

where $I_{x,y}$ is intensity of current pixel, $I_{x-1,y}$, $I_{x,y-1}$, $I_{x-1,y-1}$ are intensity of pixels from neighborhood. The new intensity include previous changes from neighborhood pixels. In result we take specific image with intensity integral properties (fig. 3).

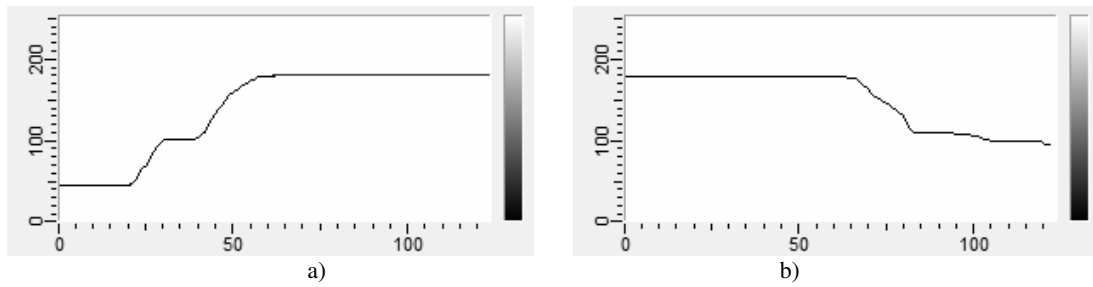


Fig. 3. Profiles of processing image of cells aggregation: a) from left to right and top to bottom direction; b) from right to left bottom and to top direction

The second stage is similar to the first but direction changed from right to left bottom and to top direction. In this case intensity is changed by low:

$$I_{x,y} = \max(I_{x,y}, I_{x+1,y}, I_{x,y+1}, I_{x+1,y+1}),$$

where $I_{x,y}$ is intensity of current pixel, $I_{x+1,y}$, $I_{x,y+1}$, $I_{x+1,y+1}$ are intensity of pixels from neighborhood.

On third stage the image are of formed as minimum from results of the first and second stages (fig. 4).

$$I_{x,y} = \min(I_{x,y}^{stage1}, I_{x,y}^{stage2}).$$

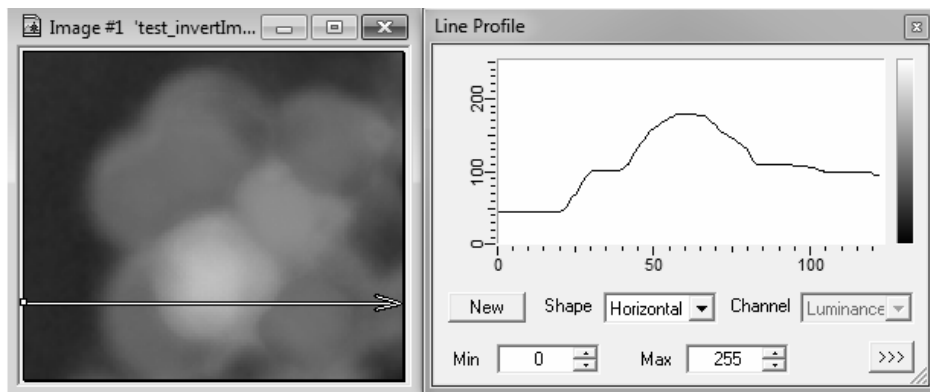


Fig. 4. Image of cells aggregation from third stage with profile line of intensity

In result we take image like as depth map that include volume properties (fig. 5).

3. Detection of cells patterns on the image

The depth map allows you to split image to layers of depth where cells are on different levels. It is realized it by multilayer binarization. We use multilayer Otsu algorithm of thresholding of image histogram (fig. 6) [2]. In result we take a few binary images with regions that are corresponding to cells of depth layers (fig. 7). The next trouble is separating such regions for cells.

The separating is very complex problem that connected with morphological properties of images [3]. We started investigate this problem from transformation of binary image to distance map [4] (fig. 8, a). Than skeleton is constructed from this map by gray thinning algorithm [5, 6] (fig. 8, a). The skeleton includes morphological features of cells aggregation regions. It consists from pixels where intensity index correspond to minimal distance from region border to this pixel. Changing of speed characterize existing circles that construct regions. Those circles correspond to cells. The special procedure spend analysis every branch of this skeleton from end points to the center and detect point of circles centers (fig. 9, a).

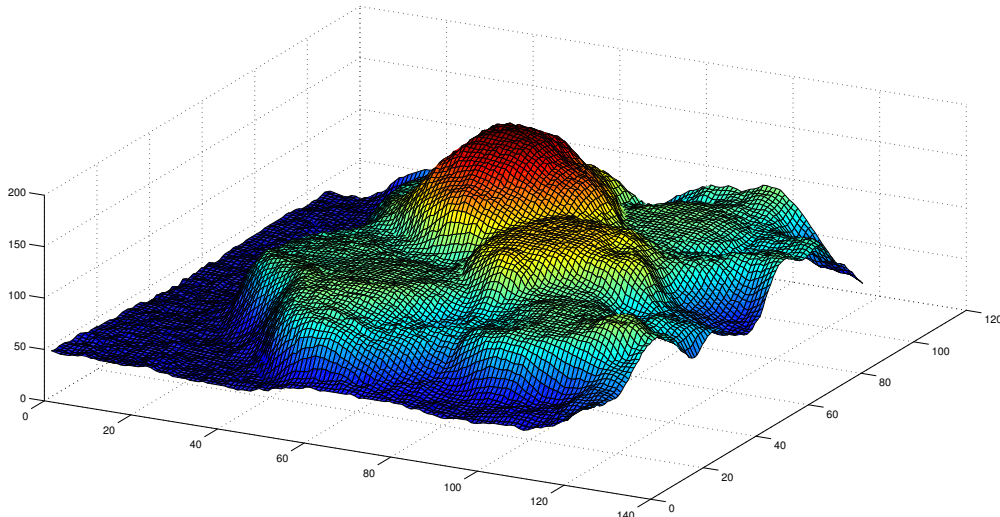


Fig. 5. Depth map of cells aggregation

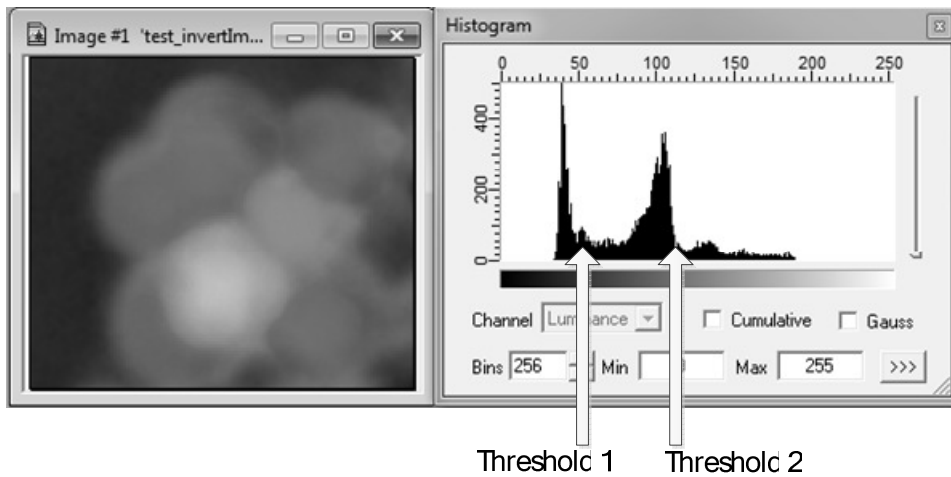


Fig. 6. Definition of binarization levels from depth map histogram

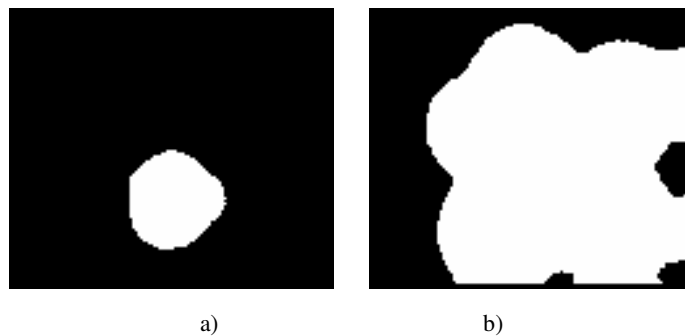


Fig. 7. Binary layers of cells aggregation: a) top; b) bottom

We detect specific pixels of center of cells where pixels index describe radius of cells (fig. 9, a). It is very easy reconstruct regions that are corresponding of separate cells by drawing circles (fig. 9, b).

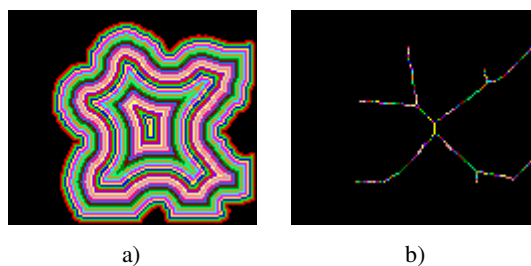


Fig. 8. Transformations of binary layers of cells aggregation: a) top; b) bottom

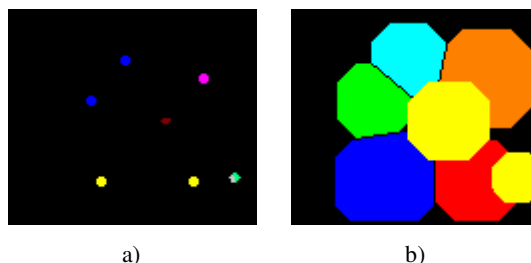


Fig. 9. Regions of separate cells from aggregation: a) centers of cells; b) reconstructed regions

Conclusion

The method of detection regions of separate living cells from aggregation is proposed. It consists of two algorithms for depth map reconstruction and cells separation. The first algorithm has complexity estimation mark which equals $O(n)$. Such mark for the second algorithm equals $O(n^2)$. It is allowed to decide the method is very fast. In other case this method leads to qualitative results. Of course shape of reconstructed cells is not ideal but this method can be improved in the near future.

References

1. Visentini-Scarzanella, M. Metric depth recovery from monocular images using Shape-from-Shading and specularities / M. Visentini-Scarzanella, D. Stoyanov, G.-Zh. Yang // Image Processing (ICIP), 19th IEEE Intern. Conference. – 2012 – P. 25–28.
2. Otsu, N. A Threshold Selection Method from Gray-level Histograms / N. Otsu // IEEE Transactions on Systems, Man and Cybernetics. – 1979. – Vol. 9, no. 1. – P. 62–66.
3. Shih, F.Y. Image Processing and Mathematical Morphology Fundamentals and Applications / F.Y. Shih. – Boca Raton : CRC Press, 2009. – P. 415.
4. Boomgaard, R. Mathematical Morphology: Extension Towards Computer Vision / R. Boomgaard. – Academisch Proefschrift. Faculteit der Wiskunde en Informatica de Universiteit van Amsterdam, 1992. – 155 p.
5. Nedzved, A. Grayscale thinning by using a pseudodistance map / A. Nedzved, S. Ablameyko, S. Uchida // Proc. of 18th Intern. Conference on Pattern Recognition ICPR. – Hong Kong, 2006. – Vol. 2. – P. 239–242.
6. Nedzved, A. Extraction of thin color pattern from images for histology investigation / A. Nedzved, V. Starovoitov // Proc. of IEEE World Congress on Computational Intelligence. – Barcelona, 2010. – P. 1922–1928.

THREE-DIMENSIONAL FAST FOURIER TRANSFORM ALGORITHM MODIFICATION BY ANALOGUE OF COOLEY-TUKEY

M. Noskov, V. Tutatchikov, I. Kolcova
Institute of Space and Information Technology Siberian Federal
University,
Krasnoyarsk, Russia
e-mail: mvnoskov@yandex.ru, vtutatchikov@mail.ru

In the article the algorithm of three dimensional fast Fourier transform, which is the three-dimensional analogue of the Cooley-Tukey algorithm, is given. The result of numerical experiments are presented.

Introduction

The one-dimensional fast Fourier transform (FFT) is the most popular tool for calculating three-dimensional Fourier transform. As a rule, to estimate the 3D FFT, a standard method of combining one-dimensional FFTs [1]. The possibility of another algorithm for 3D FFT calculation, which is an 3D analog of the Cooley-Tukey algorithm for periodic signal with a period of 2^s by three coordinate, is studied in the paper. The number of operations applied to calculate 3D FFT is considerably less than the sequential application of the 1D FFT computations. Testing of the algorithm shows that the rate of the FFT computation using the proposed algorithm is about 3 times higher than with the combination of one-dimensional FFTs.

1. The algorithm description

Let us have a look at the signal f , which is a three-dimensional periodic signal with a period of 2^s by three coordinate. The counts are given as $f_{k,t,p}$, where $k, t, p = 0 : 2^s - 1$. The discrete Fourier transform (DFT) for the signal f is given in the formula:

$$F_{l,m,n} = \sum_{k=0}^{2^s-1} \sum_{t=0}^{2^s-1} \sum_{p=0}^{2^s-1} f_{k,t,p} e^{\frac{2\pi i k p}{2^s}} e^{\frac{2\pi i m t}{2^s}} e^{\frac{2\pi i l k}{2^s}}. \quad (1)$$

The three-dimensional DFT can be calculated by means of one-dimensional DFT combinations. For this reason F is calculated as follows:

$$F_{l,m,n} = \sum_{k=0}^{2^s-1} \left(\sum_{t=0}^{2^s-1} \left[\sum_{p=0}^{2^s-1} f_{k,t,p} e^{\frac{2\pi i k p}{2^s}} \right] e^{\frac{2\pi i m t}{2^s}} \right) e^{\frac{2\pi i l k}{2^s}}. \quad (2)$$

The sum in the square brackets (2) is the one-dimensional DFT calculation, for example, in rows, then the sum in round brackets is the 1D FFT calculation in columns, then the outer sum, therefore, is the 1D FFT in height. Let us transform the formula by partition the third coordinate into even and odd components:

$$\begin{aligned} F_{l,m,n} &= \sum_{k=0}^{2^s-1} \left(\sum_{t=0}^{2^s-1} \left[\sum_{p=0}^{2^s-1} f_{k,t,p} e^{\frac{2\pi i k p}{2^s}} \right] e^{\frac{2\pi i m t}{2^s}} \right) e^{\frac{2\pi i l k}{2^s}} = \sum_{k=0}^{2^s-1} \left(\sum_{t=0}^{2^s-1} \left[\sum_{p_1=0}^{2^{s-1}-1} f_{k,t,2p_1} e^{\frac{2\pi i k 2p_1}{2^s}} \right] e^{\frac{2\pi i m t}{2^s}} \right) e^{\frac{2\pi i l k}{2^s}} + \\ &+ \sum_{k=0}^{2^s-1} \left(\sum_{t=0}^{2^s-1} \left[\sum_{p_1=0}^{2^{s-1}-1} f_{k,t,2p_1+1} e^{\frac{2\pi i k (2p_1+1)}{2^s}} \right] e^{\frac{2\pi i m t}{2^s}} \right) e^{\frac{2\pi i l k}{2^s}} = \sum_{k=0}^{2^s-1} \left(\sum_{t=0}^{2^s-1} \left[\sum_{p_1=0}^{2^{s-1}-1} f_{k,t,2p_1} e^{\frac{2\pi i k p_1}{2^{s-1}}} \right] e^{\frac{2\pi i m t}{2^s}} \right) e^{\frac{2\pi i l k}{2^s}} + \quad (3) \\ &+ \sum_{k=0}^{2^s-1} \left(\sum_{t=0}^{2^s-1} \left[\sum_{p_1=0}^{2^{s-1}-1} f_{k,t,2p_1+1} e^{\frac{2\pi i k p_1}{2^{s-1}}} e^{\frac{\pi i n}{2^s}} \right] e^{\frac{2\pi i m t}{2^s}} \right) e^{\frac{2\pi i l k}{2^s}} = g_{\alpha,\beta,0}^1(l,m,n) + e^{\frac{\pi i n}{2^s}} g_{\alpha,\beta,1}^1(l,m,n), \end{aligned}$$

where $f_{k,t,2p_1}$, $f_{k,t,2p_1+1}$ - three-dimensional even subsignals of the signal $f_{k,t,p}$, containing the signal components $f_{k,t,p}$ with even and odd indices along the third coordinate. It is clear that the dimensional of these signals is $2^s \times 2^s \times 2^{s-1}$. Observe that $g_{\alpha,\beta,0}^1(l,m,n)$ and $g_{\alpha,\beta,0}^1(l,m,n)$ is the three-dimensional FFT for subsignals $f_{k,t,2p_1}$ and $f_{k,t,2p_1+1}$ accordingly.

It can be shown that the factor $e^{\frac{\pi n}{2^s}}$ of the formula (3) is symmetric with respect to 2^{s-1} , that is for $n_1 = 0 : 2^{s-1} - 1$ and $n = n_1 + 2^{s-1}$ we have:

$$e^{\frac{\pi(2^{s-1}+n_1)}{2^s}} = e^{\frac{\pi i 2^{s-1}}{2^s}} \cdot e^{\frac{\pi i n_1}{2^s}} = e^{\frac{\pi i}{2}} \cdot e^{\frac{\pi i n_1}{2^s}} = -e^{\frac{\pi i n_1}{2^s}} \quad (4)$$

Then, from (3) and (4) we obtain:

$$\begin{aligned} F_{l,m,n_1} &= g_{\alpha,\beta,0}^1(l,m,n_1) + e^{\frac{\pi i n_1}{2^s}} g_{\alpha,\beta,1}^1(l,m,n_1), \\ F_{l,m,n_1+2^{s-1}} &= g_{\alpha,\beta,0}^1(l,m,n_1 + 2^{s-1}) - e^{\frac{\pi i n_1}{2^s}} g_{\alpha,\beta,1}^1(l,m,n_1 + 2^{s-1}), \end{aligned} \quad (5)$$

where $l, m = 0 : 2^s - 1$, $n_1 = 0 : 2^{s-1} - 1$.

For each of the sums $g_{\alpha,\beta,0}^1(l,m,n_1)$ and $g_{\alpha,\beta,1}^1(l,m,n_1)$ procedure (3) can be extended, for example, the second partition of coordinates on even and odd components. Then we obtain four sums:

$$g_{\alpha,0,0}^1(l,m_1,n_1), g_{\alpha,0,1}^1(l,m_1,n_1), g_{\alpha,1,0}^1(l,m_1',n_1), g_{\alpha,1,1}^1(l,m_1',n_1), \quad (6)$$

where $l = 0 : 2^s - 1$, $m_1, m_1', n_1, n_1' = 0 : 2^{s-1} - 1$, m_1 - even components of m , m_1' - odd ones, n_1 - even components of n , n_1' - odd ones, dimensional subsignals (6) is $2^s \times 2^{s-1} \times 2^{s-1}$.

Analogue we obtain four signals F_{l,m_1,n_1} , F_{l,m_1,n_1+2^s} , F_{l,m_1+2^s,n_1} , F_{l,m_1+2^s,n_1+2^s} .

An example, the type of signal F_{l,m_1,n_1} :

$$\begin{aligned} F_{l,m_1,n_1} &= g_{\alpha,0,0}^1(l,m_1,n_1) + e^{\frac{\pi i n_1}{2^s}} g_{\alpha,0,1}^1(l,m_1,n_1) + \\ &+ e^{\frac{\pi i m_1}{2^s}} g_{\alpha,1,0}^1(l,m_1,n_1) + e^{\frac{\pi i m_1}{2^s}} e^{\frac{\pi i n_1}{2^s}} g_{\alpha,1,1}^1(l,m_1,n_1), \end{aligned} \quad (7)$$

where $l = 0 : 2^s - 1$, $m_1, n_1 = 0 : 2^{s-1} - 1$.

Applying again on the formulas (7) procedure (3) for the decomposition of the first coordinate on even and odd components, we obtain the eight sums (8):

$$\begin{aligned} &g_{0,0,0}^1(l_1,m_1,n_1), g_{0,0,1}^1(l_1,m_1,n_1), g_{0,1,0}^1(l_1,m_1',n_1), g_{0,1,1}^1(l_1,m_1',n_1), \\ &g_{1,0,0}^1(l_1',m_1,n_1), g_{1,0,1}^1(l_1',m_1,n_1), g_{1,1,0}^1(l_1',m_1',n_1), g_{1,1,1}^1(l_1',m_1',n_1), \end{aligned} \quad (8)$$

where $l_1, l_1', m_1, m_1', n_1, n_1' = 0 : 2^{s-1} - 1$, l_1, m_1, n_1 - respectively even components of l, m, n , and l_1', m_1', n_1' - odd ones, dimensional obtained subsignals (8) is $2^{s-1} \times 2^{s-1} \times 2^{s-1}$.

Analogue we obtain eight signals F_{l_1,m_1,n_1} , $F_{l_1,m_1,n_1+2^{s-1}}$, F_{l_1,m_1+2^{s-1},n_1} , $F_{l_1,m_1+2^{s-1},n_1+2^{s-1}}$, $F_{l_1+2^{s-1},m_1,n_1}$, $F_{l_1+2^{s-1},m_1,n_1+2^{s-1}}$, $F_{l_1+2^{s-1},m_1+2^{s-1},n_1}$, $F_{l_1+2^{s-1},m_1+2^{s-1},n_1+2^{s-1}}$.

An example, the type of signal F_{l_1,m_1,n_1} :

$$\begin{aligned}
F_{l_1, m_1, n_1} = & g_{0,0,0}^1(l_1, m_1, n_1) + e^{\frac{\pi i n_1}{2^s}} g_{0,0,1}^1(l_1, m_1, n_1) + e^{\frac{\pi i m_1}{2^s}} g_{0,1,0}^1(l_1, m_1, n_1) + \\
& + e^{\frac{\pi i m_1}{2^s}} e^{\frac{\pi i n_1}{2^s}} g_{0,1,1}^1(l_1, m_1, n_1) + e^{\frac{\pi i l_1}{2^s}} g_{1,0,0}^1(l_1, m_1, n_1) + e^{\frac{\pi i l_1}{2^s}} e^{\frac{\pi i n_1}{2^s}} g_{1,0,1}^1(l_1, m_1, n_1) + \\
& + e^{\frac{\pi i l_1}{2^s}} e^{\frac{\pi i m_1}{2^s}} g_{1,1,0}^1(l_1, m_1, n_1) + e^{\frac{\pi i l_1}{2^s}} e^{\frac{\pi i m_1}{2^s}} e^{\frac{\pi i n_1}{2^s}} g_{1,1,1}^1(l_1, m_1, n_1),
\end{aligned} \tag{9}$$

where $l_1, m_1, n_1 = 0 : 2^{s-1} - 1$.

Then we can recursively apply this procedure for each of subsignals (8) until we get the final dimension subsignals $2 \times 2 \times 2$:

$$\begin{aligned}
& g_{0,0,0}^s(l_s, m_s, n_s), g_{0,0,1}^s(l_s, m_s, n_s), g_{0,1,0}^s(l_s, m_s, n_s), g_{0,1,1}^s(l_s, m_s, n_s), \\
& g_{1,0,0}^s(l_s, m_s, n_s), g_{1,0,1}^s(l_s, m_s, n_s), g_{1,1,0}^s(l_s, m_s, n_s), g_{1,1,1}^s(l_s, m_s, n_s),
\end{aligned} \tag{10}$$

where $l_s, l'_s, m_s, m'_s, n_s, n'_s = 0 : 1$, l_s, m_s, n_s - respectively even components of $l_{s-1}, m_{s-1}, n_{s-1}$, and l'_s, m'_s, n'_s - odd ones, dimensional obtained subsignals (10) $2 \times 2 \times 2$. The final values for F are calculated similarly to (9).

Let us calculate the total amount of operations. The three-dimensional FFT by analogue of the Cooley-Tukey algorithm over the signal $f_{k,t,p}$ of the $2^s \times 2^s \times 2^s$ dimensional requires $3s2^{3s}$ operations of complex multiplications and $\frac{7}{8}s2^{3s}$ operations of complex additions [2, 3]. If we calculate the 3D FFT of the ignition signal $f_{k,t,p}$ by partition into 1D FFT, the number of operations required will be: $s2^{3s}$ of complex multiplication and $s2^{3s}$ of complex additions.

2. The obtained results

For the algorithm testing the program in the programming language C++ has been written. The testing was conducted on PC with following characteristics:

Processor: Intel Core i5 2.4 GHz;

RAM: 4 GB;

Operating system: Windows 7.

The presented algorithms working time is given in table.

Table 1

The three-dimensional FFT working time, sec.

Size of signal	3D FFT by partition into 1D FFT	3D FFT of analogue algorithm of Cooley-Tokey
32*32*32	0,006	0,002
64*64*64	0,045	0,019
128*128*128	0,298	0,134
256*256*256	2.850	1,225
512*512*512	29,840	10,658

Conclusion

The modified algorithm of the three-dimensional FFT by analogue of the Cooley-Tukey algorithms requires fewer complex operations of multiplication and is faster than its analogue of 3D FFT calculation by partition into 1D FFT.

References

1. Pissis, S. Parallel Fourier Transformations using shared memory nodes / S. Pissis // MSc in High Performance Computing. – The University of Edinburgh, 2008.

2. Starovoytov, A.V. About multidimensional analog of algorithm of Cooley-Tukey / A.V. Starovoytov // The bulletin of the Siberian state space university of a name of academician M.F. Reshetnev. – Iss. 1 (27). – 2010. – P. 69–73. [In Russian]
3. Tutatchikov, V.S. Calculating the n-Dimensional Fast Fourier Transform / V.S. Tutatchikov, O.I. Kiselev, M.V. Noskov // Pattern Recognition and Image Analysis. – 2013. – Vol. 23, no. 3. – P. 429–433.

DEVELOPMENT OF ALGORITHM FOR LOCALIZATION OF SYMBOLIC WITH THE USE OF ANALYSIS OF THE COLOR MAP

A.A. Orlov, A.V. Astafiev, D.P. Popov
Murom Institute (branch) of Vladimir State University
named after A.G. and N.G. Stoletovs, Russia
e-mail: AlexeyAlexOrlov@gmail.com

The article shows relevance of design and implementation of systems of localization of production markings on digital images. The analysis, description and comparison of this class of systems is performed. In the course of conducting a comprehensive analysis the main characteristics of the systems under consideration are identified and existing software systems are analyzed. The results of estimates of the characteristics of the systems on the example of localization systems of state license plates are given. The algorithm for localization of production markings with the use of analysis of the color data on digital images is developed.

Introduction

Technical (machine) vision systems (TVS) [1] allowing to recognize objects and control quality of products are becoming more common in industrial automation. The latest technology allows creation of cameras similar in structure to the human eye. Intelligent machine vision cameras are used in many fields of industry and science. For the purpose of automation of production processes there are problems of selecting and creating algorithms for searching, localization and identification of industrial objects [2]. One of such problems is the problem of search and localization of labels on the monitored objects, because for the correct operation of the TVS it is necessary to find the image that gives you the maximum amount of useful information [3]. Contributions to the development of techniques for the identification of markers in TVS were made by Alan Sloan, Georey M. Davis, George H. Freeman, Huawu Lin and many others [4].

The aim of the research is the development of an algorithm for localization of production markings using the analysis of color digital images.

To achieve this goal it is necessary to perform several tasks:

Review existing localization systems and identify the main functional features of these systems.

Review existing algorithms for localization of production markings

Form the main requirements for the developed algorithm of localization of production markings on the basis of the analysis of color data.

Perform experimental research of the algorithm.

1. Comparative analysis of localization systems marking on digital images

At the first stage of the analysis several systems of localization of state license plates of cars in digital images and video streams created by leading domestic and foreign manufacturers [5] were considered.

The “AUTO-Inspector” system. AUTO-Inspector, which is a special hardware-software module for registration and recognition of license plates made by the “STB-Service” company, works as follows. The TV camera is installed over the controlled lane. The camera is connected to the AUTO-Inspector system, the software of which detects the appearance of the car in a frame, select the frame with the optimum size and clarity of the car license plate and recognizes the number of the car in the frame. The database retains the freeze frame or the entire movie associated with the vehicle, as well as the recognized license plate number, date and time of travel of the vehicle. The result is a string with the recognized number and freeze frame with a picture of the car sent to the database.

The “Intelligent Auto” system. The ITV company has developed a system of recognition of automobile numbers called “Intelligent Auto”. The principle of action of the system of recognition of automobile numbers is as follows: when driving on a stretch of roadway, in the field of vision of the camera, automatic recognition of license plate is performed, the license plate number is logged and checked for a match with numbers in the database.

The “SL-Traffic” system. SL-Traffic is a software module for reading and recognition of automobile numbers in real time made by the SpecLAB company. The module works as a video filter for the GOALcity system or can be integrated into any video registration system, including third-party systems. The principle of action is as follows: At the checkpoint a camcorder is set according to the

requirements listed above. The video signal from the camera is entered into a computer by means of installed video capture boards and processed in real time by the SL-Traffic software module for reading and recognition of automobile numbers.

The review highlighted important criterias to evaluate the systems of localization of state license plates of cars in digital images and video streams:

- the maximum permissible speed (in kilometers per hour);
- the maximum number of simultaneously recognized license plate numbers;
- minimum illumination (in Lux);
- vertical angle to the plane of the roadway (in degrees);
- the probability of localization and recognition (in percentage);
- minimum height of the number in the frame (in pixels).

2. Comparative analysis of algorithms for the localization of character markings on digital images

The Scale Invariant Feature Transform (SIFT) algorithm. This algorithm belongs to the methods to compare images based on the comparison of knowledge about images in general. In the general case, it looks like the following: for each point of the image the value of a specific function is calculated, on the basis of these values a certain characteristic can be attributed to the image, then the task of image comparison is reduced to the comparison of such characteristics.

The algorithm can be represented by a sequence of the following steps:

1. Key points and their descriptors are detected on digital images.
2. Matching key points are detected by matching descriptors.
3. Localization of the specified object is performed based on a set of matching key points.

The Speeded Up Robust Features (SURF) algorithm. The SURF algorithm solves two tasks – finding key points of the image and creating their descriptors that are invariant to the scale and rotation. A key point descriptor will be the same even if the sample was resized and rotated. The method locates special points using the Hesse matrix. The determinant of the Hesse matrix reaches an extremum at points of the maximum change in the gradient of brightness. This algorithm well detects spots, corners and edges of the lines. The determinant of the Hesse matrix is invariant with respect to rotation, but not invariant to scale. The SURF algorithm uses different filters to find the determinant of the Hesse in order to solve the problem of scale invariance.

For each key point the direction of maximum brightness change (gradient) and the scale taken from the scale coefficient of the Hesse matrix are calculated. Gradient at the point is calculated using Haar filters.

After finding the key points SURF forms their descriptors. A descriptors is a set of 64 (or 128) numbers for each key point. These numbers reflect fluctuation of the gradient around the key point. Since the key point is the maximum of the Hessian, this ensures that in the neighborhood of the point there are areas with different gradients. Therefore there is high dispersion (difference) between descriptors for different key points.

3. Algorithm for localization of production markings with the use of analysis of the color data on digital images

Development of the algorithm is based on the fact that each code used for production purposes has its own color palette. For example, the state license plate numbers of automobiles are using white and black colors, as are markings based on bar codes (fig. 1) [5].



Fig. 5. Types of labels

Therefore it is supposed that during the analysis of production images the method of comparison of the color data of the input digital image with the reference will be used [6].

Designed localization algorithm can be represented as a sequence of the following actions:

In accordance with the size of the input image the matrix of deviations of color parameters from the reference is formed using the formula 1 [7]:

$$C_{i,j} = |\text{Color} - M_{i,j}|, \quad (1)$$

where color – the reference color described by the RGB vector; $C_{i,j}$ – color of the current pixel of the image described by the RGB vector.

Values of matrix are calculated for each line of the image with the cumulative total according to the formula 2:

$$C_{L,i} = C_{L,i} + C_{L,i-1}, \quad (2)$$

Each line of the image is smoothed using the function 3;

$$C_{L,i} = \begin{cases} C_{i,j+L} - C_{i,j}, & \text{if } j < W - L \\ C_{i,j}, & \text{if } j > W - L \end{cases} \quad (3)$$

where W is the width of the input image.

The maximum value from the matrix C is found and the matrix of probabilities of finding a label at the given point is calculated using the formula 4;

$$C_{L,i} = \frac{\max - C_{L,i}}{\max}, \quad (4)$$

where \max – the maximum value of the C matrix.

The algorithm output is the matrix where each element shows with what degree the analyzed image pixel belongs to the field of marking. An example of the use of the algorithm is shown in fig. 2.



Fig. 6. Example of use of the algorithm

Thus, knowing the approximate aspect ratio of the production marking, the one area that relates to the field of the marking could be selected from the resulting three areas (fig. 3).



Fig. 7. The selection marker by aspect ratio

4. Experimental researches of the developed algorithm

Localization was performed using two algorithms: the developed algorithm and the SURF algorithm.

The reliability was identified as 96% for the developed algorithm, and 92% for the SURF algorithm. The obtained reliability of the algorithm is due to the fact that the images contain objects similar in color to the reference. Low reliability of the SURF algorithm is due to the fact that a large number of labels are not similar to each other in form and content. The average speed of the algorithm is 1.2 seconds for the image with a resolution of 640x480 pixels.

For the experimental research examples of digital images from real production processes were taken (fig. 4).



Fig. 8. The results of the work of the algorithm

The developed algorithm fully meets the current requirements. Reliability of the method is 96%. The average speed of a method is 1.2 seconds for the image with a resolution of 640x480 pixels.

References

1. Orlov, A.V. Implementation and application of the algorithm of digital localization images character tags on the basis of the analysis of the rate of change of brightness / A.V. Orlov, A.V. Astafiev // Modern problems of science and education. – 2012. – № 6.
2. Kovalenko, E.N. The methods for the license plate on the image [Electronic resource] / E.N. Kovalenko, A.V. Sytnik // Scientific-research Institute of applied electronics. – Mode of access : <http://visa.net.ua/content/maket015.pdf>.
3. Gaganov, K.A. Segmentation of moving objects in video stream / K.A. Gaganov // Scientific-popular online magazine Graphics and Multimedia.
4. Gashnikov, M.V. The methods of computer processing of images / M.V. Gashnikov, N.I. Glumov, N. Ilyasov. – M. : FizMatLit, 2003. – 784 p.
5. Gonzalez, P. Digital processing of images / P. Gonzalez. – M. : Technosphere, 2005. – 1072 p.
6. Orlov, A.V. Analysis of visual systems of monitoring of the production process at the industrial enterprises / A.V. Orlov, A.V. Astafiev, A.V. Provotorov // Scientific “Journal of NGYIY“. – 2011. – No. 1, vol. 114. – P. 26–32.
7. Orlov, A.V. Development of a technique for the automatic identification of industrial products on the basis of the analysis of methods of marking / A.V. Orlov, A.V. Astafiev, A.V. Provotorov // Engineering: problems and perspectives. – 2011. – No. 1.

RESEARCH METHODS FOR CLASSIFICATION OF THE CRYSTALLOGRAMS IMAGES

R.A. Paringer¹, A.V. Kupriyanov²

¹S.P. Korolyov Samara State Aerospace University
(National Research University), Russia;

²Image Processing Systems Institute of the RAS, Samara
e-mail: RusParinger@gmail.com

This paper analyses the possibility of the crystallogram images classification, based on the geometric characteristics. Classification is based on SVM method. For estimation of features of the crystallograms images follow methods were used: estimating geometrical features of dendritic crystallograms and calculating the form factors of the spatial spectrum of the images.

Introduction

Image analysis is an important part of crystallograms task in medical diagnostics. Automating the process of crystallograms processing will improve the quality of diagnosis and reduce the time required. Crystallogram correspond to a certain type of characteristic values of diagnostic features. It is necessary to select the features that allow us to determine the type of a particular image crystallogram more accurately. There are many different methods for analysis the dendrite crystallogram images [1–3]. The estimating geometrical features and method for calculating form factors were selected. For the classification Support Vector Machine method is selected.

1. Geometrical features

The dendrite model is presented on fig.1. A, B, C, D are «top» key elements, E, F, G are «root» key elements, EF, FG are distances between branches. EF with FG is a stem of dendrite, AE, BE, CF, DG are dendrites' branches.

The method consists in the following sequence of actions: threshold and median filters [4], skeletonization, discrimination of key elements, building a «map of dendrites», calculation of geometrical features [5].

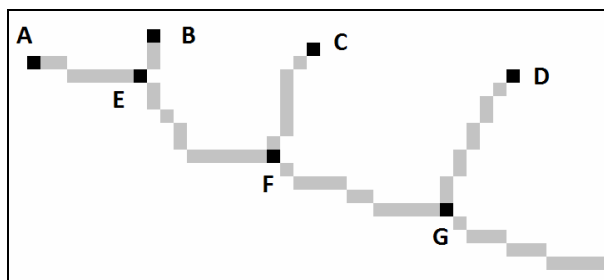


Fig. 1. Dendrite model

«Growth» factor F_g is the ratio of the lengths of all branches to the sum of the distances between them. For model calculated as follow:

$$F_g = \frac{AE + BE + CF + DG}{EF + FG}. \quad (1)$$

«Angle» factor F_a is the ratio of the sum of the angles of all branches the total number of them.

«Symmetry» factor F_s is the ratio of the number of «top» key elements to the number of «root» key elements. For model $F_s = 4/3$.

2. Form factors

The information contained in the crystallogram is structurally redundant. Visually, these images are perceived as a set of contour lines, subject to certain rather complex sequence. If we consider the spatial spectrum of such images, it would be located in a sufficiently narrow frequency band in a characteristic spatial frequency, which can be called carrier frequency. It is known that if in the original image collinear stripes of a certain direction prevail, then in Fourier transform of the original image stripes of the same direction will prevail.

Threshold the image spectrum can clearly distinguish the shape of the spectrum and calculate the form factors.

Consider the various geometric features of the spectral shape. This group includes those features, the calculation of which is based on the geometric characteristics appearing on the image objects. This features were called form factors. To calculate the form factors, it is necessary to calculate a perimeter and an area of the object. If we take an image pixel per unit area, then the area of the object is equal to the number of pixels. Calculation of the perimeter is not a trivial task and requires a specific approach.

The general algorithm can be represented by the following sequence of actions: calculation of the image spectrum, threshold processing, transformation of the domain into a closed domain, extraction of the object contour, calculation of the form factors [6].

S is the number of pixels in an image of a closed region of the spectrum. P is the number of pixels equal to the perimeter of the closed contour of the spectrum.

«Blair-Bliss» coefficient F_b :

$$F_b = \frac{S}{\sqrt{2 \cdot \pi \cdot \sum_i r_i^2}}, \quad (2)$$

where i is the number of object's pixel;

r_i is the distance from the object's pixel to the object's center of gravity.

«Malinowska» coefficient F_m :

$$F_m = \frac{P}{2 \cdot \sqrt{\pi \cdot S}} - 1. \quad (3)$$

«Haralick» coefficient F_h :

$$F_h = \sqrt{\frac{(\sum_i d_i)^2}{n \cdot \sum_i d_i^2 - 1}}, \quad (4)$$

where d_i is the distance of the outline's pixels from the object's center of gravity;

i is the number of outline's pixel;

n is the number of objective's outline pixels.

«Compact» coefficient F_c :

$$F_c = \sqrt{\frac{1}{N} \sum_i \left(r_i - \frac{1}{N} \sum_i r_i \right)^2}, \quad (5)$$

where N is the total number of objective's pixels;

i is the number of object's pixels;

r_i is the distance of the object's pixels from the object's center of gravity.

3. Classification

(This) Classification (is) based on Support Vectors Machine method.

For the classifier training jpg format images with a resolution of 256 by 256 pixels, consisting of 256 grayscale tones, were used. A total of 1187 crystallograms images were processed. These images belong to two different classes.

The first class includes dendritic crystallograms (fig. 2). There are 468 images altogether.



Fig. 2. Dendrite crystallogram

The second class includes radiant crystallograms (fig. 3). There are 719 images in total.



Fig. 3. Radiant crystallogram

The experiment was performed at fixed values of the parameters for the crystallograms of all classes. Estimation of the geometrical features was carried out using an adaptive threshold filter with a size of the window equal to 64 and a median filtering parameter equal to 2. A threshold with parameter value equal to 20 performs the estimation of spectral features. The values were chosen experimentally.

Let us define the quality criterion of classification through the classification error, showing the quantity of cases, in which the classifier has made a wrong decision, as follows:

$$\varepsilon = \frac{k}{n} \cdot 100\%, \quad (6)$$

where k is the number of classification errors;

n is the total number of images.

An investigation of the classification error was performed.

Table 1 shows the values of the classification error for different combinations of features. In the columns of features sign «+» means that the feature is taken into account in the experiment, sign «-» means that the feature is not taken into account.

The value of misclassification of selected coefficients

№	F_g	F_a	F_s	F_m	F_c	ε
1	-	+	+	+	+	1.52
2	+	+	+	+	+	1.52
3	-	+	+	+	-	1.77
4	+	+	+	+	-	1.77

Conclusion

As results of the research, the most characteristic features of the dendritic crystallograms images were selected. They include “Growth”, “Angle”, “Symmetry”, “Malinowska” and “Compact”.

This work was finically support of the RFBR grant (#12-01-00237-a), the ONIT RAS program #6 “Bioinformatics, modern information technologies and mathematical methods in medicine” 2012, the state task from the Ministry of Education and Science of the Russian Federation 2012-2013 (# 8.3195.2011).

References

1. Ilyasova, N.Yu. Analysis of Features of Texture Images for Crystallogram Identification and Classification / N.Yu. Ilyasova, A.V. Kupriyanov, A.G. Khramov // Optical Memory & Neural Networks. – 2002. – Vol. 11, no. 1. – P. 19–28.
2. Kupriyanov, A.V. Ophthalmic Pathology Diagnostics Using Textural Features of the Lachrymal Fluid Crystal Images / A.V. Kupriyanov, N.Yu. Ilyasova, A.G. Khramov // Pattern Recognition and Image Analysis (Advances in Mathematical Theory and Applications). – 2005. –Vol. 15, no. 4. – P. 657–660.
3. Kupriyanov, A.V. Statistical Features of Image Texture for Crystallogram Classification / A.V. Kupriyanov, A.G. Khramov, N.Yu. Ilyasova // Pattern Recognition and Image Analysis. – 2001. – Vol. 11, no. 1. – P. 180–183.
4. Soifer, V.A. Computer image processing. Part 2. Methods and algorithms / V.A. Soifer // VDM Verlag Dr. Müller. – 2010. – February 18. – P. 584.
5. Paringer, R.A. Methods For Estimating Geometric Parameters of The Dendrite’s Crystallograms / R.A. Paringer, A.V. Kupriyanov // Proceedings of 8th Open German-Russian Workshop “Pattern Recognition and Image Understanding” OGRW-8-11. – 2011. – P. 226–229.
6. Ilyasova, N.Yu. Information technologies of image analysis for purposes of medical diagnosis / N.Yu. Ilyasova, A.V. Kupriyanov, A.G. Khramov. – M. : Radio and communication, 2012. – 424 p. [In Russian]

LOGICAL-AND-PROBABILITY SIMULATION MODEL OF DATA PROCESSING AND ANALYSIS

G. Polyakova
Novosibirsk State University, Russia
e-mail: Karusel2007@ya.ru

In this paper the method of analysis of empirical information represented in the form of multidimensional time series is suggested. These time series reflect astrophysical, climatic features and tick-born encephalitis disease's indices. Results of statistical analysis are represented in the form of logical rules reflected cause-effect relations of object under investigation.

Introduction

Methods to build the logical-and-probabilistic models for the analysis of multivariate time series are proposed in the works [1, 5–7].

The paper considers the construction of logical-and-probability models as the form of a decision tree using a search of optimal division of multidimensional variable space. The forecasting method using the class of logical decision functions has some distinctive features: data only contain a limited number of time points, and results must be easy to interpret. A search of optimal division of variable space is based on sort out similar to branch and bound method. It shows that logical regularities are looking at statements with high probabilistic abilities. The regularity list corresponds to a partitioning of the initial space of variables on a number of areas; each one corresponds to the true of regularity from the list. These studies use approaches, applied earlier for solving the similar statistical problems [5, 6]. The forecasting areas are especially powerful because as few as one future sample can be used in determining the prediction intervals. The events, really occurring in the nature often have complex description, containing heterogeneous (quantitative, as well as qualitative) information. Effectiveness of the suggested method is shown by solving applied problems in the sphere of ecology.

1. A method of obtaining logical regularities

Let Γ be a population of elements or objects under investigation, and let v be a sample of observed objects from population Γ . Suppose $X(a)$ is a description of observed object $a \in \Gamma$, and $Y(a)$ is a value of purposeful quantitative variable Y . We assume that $X(a) = (X_1(a), \dots, X_j(a), \dots, X_n(a))$, $Y(a) = (Y_1(a), \dots, Y_j(a), \dots, Y_m(a))$, where the sets X and Y may simultaneously contain qualitative and quantitative features X_j , $j = \overline{1, n}$; or Y_j , $j = \overline{1, m}$; respectively. Let D_j be the domain of the variable X_j , $j = \overline{1, n}$; and D_y the domain of the variable Y , $D_y = [\alpha, \beta] \subset \mathbb{R}$. The feature space is given by the product set $D = \prod_{j=1}^n D_j$.

By W_j denote sets of subsets $E_j \subseteq D_j$, such that $E_j = [\alpha_j, \beta_j]$ if X_j is a quantitative or ranged variable, and E_j is a finite subset of values of variable if X_j is a qualitative variable with non-ranged values; $j = \overline{1, n}$.

We shall say that a set E is a rectangular set in D if

$$E = \prod_{j=1}^n E_j, \text{ where } E_j \in W_j.$$

Denote by $J(a, E_j)$ an one-place predicate « $X_j(a) \in E_j$ », where $a \in \Gamma$, $E_j \in W_j$.

A conjunction $S(a, E) = J(a, E_{j_1}) \wedge \dots \wedge J(a, E_{j_m})$ is called a conjunction of length m .

Fix some interval $I_i \subseteq D_y$. Let $N(I_i)$ be a quantity of objects from the sample v such that $Y(a) \in I_i$.

For any conjunction $S(a, E)$ by $N(I_k, S)$ denote a quantity of objects from the sample v such that $Y(a) \in I_k$ and conjunction S is true, if $I_k = \max I_i$, and by $N(I_i, S)$ denote a quantity of objects such that $Y(a) \in (D_y \setminus I_k)$ and conjunction S is false.

A conjunction $S(a, E)$ is a logical regularity if

$$\frac{N(I_k, S)}{N(I_k)} \geq \delta, \quad \frac{N(I_i, S)}{N(I_i)} \leq \beta, \quad (1)$$

where δ and β are thresholds decided by the user; $0 \leq \beta < \delta \leq 1$. Greater δ and smaller β correspond to stronger logical regularities. Denote by S^* the set of logical regularities.

A conjunction $S(a, E)$ is a potential logical regularity

$$\frac{N(I_k, S)}{N(I_k)} \geq \delta, \quad \frac{N(I_i, S)}{N(I_i)} > \beta. \quad (2)$$

Denote by S' the set of potential logical regularities. Note that may exist S^* one-place predicates $S' \wedge J(a, E_j) \wedge \dots$

If for some conjunction $S(a, E)$ we have

$$\frac{N(I_k, S)}{N(I_k)} < \delta \quad (3)$$

then the conjunction S could not be a logical regularity (even if add any one-place predicates). Denote by S the set of such conjunctions. Obviously, any conjunction $S(a, E)$ belongs to one of the sets S^* , S' , S .

Obtaining logical regularities

In this paper, the sets W_j are formed on base of values of the variables X_j in a sample v . It can be assumed that in real data the values $|W_j|$ are low enough.

In order to all non-redundant logical regularities to be obtained from multidimensional sample v , the following steps are performed.

– Step 1. Consider all possible conjunctions of the length 1, i.e. conjunctions of the type $S(a, E) = J(a, E_j)$, where $E_j \in W_j$, $j = \overline{1, n}$. If conjunction $S(a, E)$ satisfies condition (1), then the statement $S(a, E)$ is included in the list of logical regularities S^* , and corresponding subset E_j is excluded from further consideration. If conjunction $S(a, E)$ satisfies condition (3), then corresponding subset E_j is excluded from further consideration. Denote by W_j^1 the set of subsets E_j kept for further analysis;

– Step 2. Consider all possible conjunctions of the length 2, i.e. for $j \neq l$ conjunctions of the type $S(a, E) = J(a, E_j) \wedge J(a, E_l)$, where $E_j \in W_j^1$, $E_l \in W_l^1$. If conjunction $S(a, E)$ satisfies condition (1), then the statement $S(a, E)$ is included in the list of logical regularities S^* . If conjunction $S(a, E)$ satisfies condition (2), then the pair $\langle E_j; E_l \rangle$ is included in the temporary list of potential regularities. Denote by W_j^2 the set of pairs $\langle E_j; E_l \rangle$ in the list;

– Step 3. Consider all possible conjunctions of the length 3, i.e. for $j \neq l \neq m$, $j \neq m$ conjunctions of the type $S(a, E) = J(a, E_j) \wedge J(a, E_l) \wedge J(a, E_m)$, where $E_j \in W_j^2$, $E_l \in W_l^2$, $E_m \in W_m^2$, such that all possible pairs $\langle E_j; E_l \rangle$, $\langle E_j; E_m \rangle$, and $\langle E_l; E_m \rangle$ belong to the list. If conjunction $S(a, E)$ satisfies condition (1), then the statement $S(a, E)$ is included in the list of logical regularities S^* . If conjunction $S(a, E)$ satisfies condition (2), then the trinity $\langle E_j; E_l; E_m \rangle$ is included in the temporary list of potential regularities. Denote the set of trinities $\langle E_j; E_l; E_m \rangle$, and so on.

It can be assumed that in real data quantities of elements of sets W_j^m are quickly decrease as number m of steps is increases. Notice that in some data for lesser β (near 0) and greater δ (near 1) there is no logical regularity at all. On the other hand, for fixed small β (for example, $\beta = 0,05$), if δ is decreased ($\delta \rightarrow \beta$), then a number of logical regularities may quickly increase with reduction of their quality. Therefore, for fixed β we must choose δ with a few regularities. It can be obtained by successive decrease of δ with certain step $\Delta\delta$.

2. Applications

The proposed methods were tested on artificial and real data. For example, they were applied in analysis of the present situation for tick-borne encephalitis. Tick-borne encephalitis (TBE) is the most important flavivirus infection in Europe, Russia, and the Far East [2- 4, 8-10]. The risk for contracting TBE depends strongly on the density of the infected questing ticks and many studies have investigated tick population dynamics and the parameters affecting them. To understand the fluctuation in incidence rates during the period of observation a complex interrelation of several factors has to be considered, such as astrophysical and climatic factors. It can be assumed that changes of leisure activities in nature, increasing/decreasing mobility to risk areas, changes in wildlife hosts/tick populations may have influenced the quantity and quality of epidemiological data. On the other hand, it's a known fact that changes of solar activity have influence on wildlife populations, see, for instance [9]. So, in the paper [9] studies have shown a clear influence of climatic and landscape arrangements in the microhabitat on tick abundance and dynamics. As in [3], locations were selected in endemic foci of TBE with different intensity, which were identified according to the presence of human disease. The author of this paper assessed the influence of climatic factors on questing activity of ticks over a 22-year period at 3 locations in Russia (Novosibirsk, Gorno-Altaysk and Irkutsk). Sites differ according to various factors, such as climate, amount of rain, height above sea level, vegetation and wildlife, all three stages of ticks were collected monthly. The purpose of this study was to relate observed differences in ticks questing activity to local climate.

The proposed methods were used, in particular, to obtain cause-effect relations between climatic factors (air temperature (t), humidity (h), precipitation and dew point), astrophysical factors (solar radiations per unit of horizontal (s) and perpendicular surface, and reflected radiation), and TBE indices in Novosibirsk region of Russia [7]. Data of the last 22 years (1991 — 2012) were analyzed. Purposeful variable Y is an annual number of the TBE incidences per 100000 people. Let us introduce some constructed regularities for Novosibirsk:

– if ($t4 \in [-1,5; 3]$ & $t7 \in [19; 22,2]$) or ($t7 \in [19; 22,2]$ & $t8 \in [15; 18,7]$ & $t10 \in [2;1; 5,6]$), then $Y(a) \in [20; 62]$, where $t4; t7; t8; t10$ are average air temperature for April, July, August, and October;

– if ($h3 \in [71,3; 78]$ & $h8 \in [68; 74,5]$) or ($h6 \in [64; 73,6]$ & $h8 \in [68; 74,5]$ & $h10 \in [73,5; 80]$), then $Y(a) \in [2; 19]$, where $h3; h6; h8; h10$ are average humidity for March, June, August, and October;

– if ($s5 \in [1400; 2300]$ & $s7 \in [1611; 2300]$) or ($s4 \in [285; 2070]$ & $s7 \in [1611; 2300]$ & $s9 \in [470; 1670]$), then $Y(a) \in [2; 19]$, where $s4; s5; s7; s9$ are solar radiations per unit of horizontal surface for April, May, July, and September;

– if ($w5 \in [30; 90]$ & $w7 \in [30; 135]$) or ($w2 \in [55; 135]$ & $w3 \in [40; 85]$ & $w12 \in [50; 110]$), then $Y(a) \in [20; 62]$, where $w2; w3; w5; w7; w12$ are Wolf numbers for February, March, May, July, and December.

Joint analysis of multidimensional time series of annual tick-borne encephalitis (TBE) prevalence among Novosibirsk, Irkutsk and Gorno-Altaysk populations, monthly air temperatures, relative humidity, rainfalls, annual solar activity expressed as Wolf's numbers and vaccination rate was performed using logical-and-probabilistic models. In order to determine the general relevant factors, but did not rule out the specifics of each region, was performed joint information processing on all the regions. Analysis of the influence of natural factors of the current year (last 12 months) for the value of Y , and the Y forecast in the current year were carried. When processing the raw data for the variable Y was selected three intervals, there are three images. The first image was determined by the size of the TBE incidence up to 10 (low incidence) and the second from 10 to 20 (average), the third is greater than 20 (high). A tree solutions was built by partitioning the space variables with the help of algorithm. Obtained solution was optimum by criterion of minimizing the risk of erroneous classification.

Let $v_1^t = (T_{10}^{t-1}, T_{11}^{t-1}, T_{12}^{t-1}, T_1^t, \dots, T_9^t)$ – data on average monthly air temperatures ($^{\circ}\text{C}$) the relevant year t ; $v_2^t = (V_{10}^{t-1}, V_{11}^{t-1}, V_{12}^{t-1}, V_1^t, \dots, V_9^t)$ and $v_3^t = (O_{10}^{t-1}, O_{11}^{t-1}, O_{12}^{t-1}, O_1^t, \dots, O_9^t)$ – data on average monthly humidity (%) and precipitation respectively (mm); $X^t = (T_1^t, \dots, T_{12}^t, V_1^t, \dots, V_{12}^t, O_1^t, \dots, O_{12}^t)$ – data on natural factors in year t , where the subscripts denote the serial number of the month. Training picks for the current year is $v = \{v_1^t, v_2^t, v_3^t, Y^t\}$, and in the background three years – $v = \{X^{t-2}, X^{t-1}, X^t, Y^{t+1}\}$. Let characteristics of parasitic system TBE: the number of mammals-host (Z^t); the proportion of affected animals (IE^t - larvae and nymphs, IEl^t - larvae, $IE n^t$ - nymphs); the index of abundance of parasites on animals (IO^t - larvae and nymphs, IOl^t - larvae, $IO n^t$ - nymphs); number of ticks (Nk^t), where the subscript i specifies the month ordinal of a year t ; top index t - the number of the year. The logical-probabilistic models received the following patterns for current forecast (let P_i^* Bayesian estimation of the probability of correct classification for the i patterns obtained on formula [6, p. 96]):

– if $(V_{11}^{t-1} \in [68,5\%, 80,6\%])$ and $(V_4^t \in [64,5\%, 75\%])$ and $(V_6^t \in [42,3\%, 77\%])$ and $(T_6^t \in [13^{\circ}\text{C}, 17^{\circ}\text{C}])$, then $Y^t = 3$; $P_1^* = 0,75$; IF $(V_{11}^{t-1} \in [68,5\%, 80,6\%])$ and $(V_6^t \in [66,5\%, 77\%])$ and $(T_6^t \in [17^{\circ}\text{C}, 21^{\circ}\text{C}])$, then $Y^t = 2$; $P_2^* = 0,6$.

– if $(V_{11}^{t-1} \in [68,5\%, 80,6\%])$ and $(V_4^t \in [46\%, 64,5\%])$ and $(V_6^t \in [66,5\%, 77\%])$ and $(T_6^t \in [13^{\circ}\text{C}, 17^{\circ}\text{C}])$, then $Y^t = 1$; $P_3^* = 0,75$; IF $(V_{11}^{t-1} \in [80,5\%, 89\%])$ and $(V_6^t \in [42,3, 66,5\%])$ and $(T_6^t \in [17^{\circ}\text{C}, 21^{\circ}\text{C}])$, then $Y^t = 1$; $P_4^* = 0,75$.

To search for patterns for the target variable (Z^t or Nk^t), the range of values Z^t and Nk^t were selected three intervals, there are three images. To predict the target variable natural factors were considered (selected on the basis of the results on the incidence of TBE): temperature of April of the current year, previous and 2 years ago ($T_4^t, T_4^{t-1}, T_4^{t-2}$); temperature in November of the previous year, 2 and 3 years ago ($T_{11}^{t-1}, T_{11}^{t-2}, T_{11}^{t-3}$), humidity November of the previous year, 2 and 3 years ago ($V_{11}^{t-1}, V_{11}^{t-2}, V_{11}^{t-3}$); as well as data on parasitological system TBE for previous years: $IE^{t-1}, IE^{t-2}, IE^{t-3}$; $IO^{t-1}, IO^{t-2}, IO^{t-3}$; Nk^{t-1}, Nk^{t-2} . At a temperature of April of the current year from 6.35°C is the number of ticks Nk^t in this year above the average for the last 20 years (more than 18 ticks at flago-km).

[IF $(-1.50 \leq T_4^t < 6.35)$ and $(13.45 \leq Nk^{t-1} < 51.80)$ and $(1.67 \leq IO^{t-1} < 6.09)$ and $(-5.50 \leq T_{11}^{t-3} < -2.60)$] or [IF $(-1.50 \leq T_4^t < 6.35)$ and $(13.45 \leq Nk^{t-1} < 51.80)$ and $(32.24 \leq IE^{t-1} < 66.02)$ and $(-8.85 \leq T_{11}^{t-3} < -1.00)$] or [IF $(-1.50 \leq T_4^t < 6.35)$ and $(13.45 \leq Nk^{t-1} < 51.80)$ and $(32.24 \leq IE^{t-1} < 66.02)$ and $(3.1 \leq IO^{t-1} < 6.09)$], then in the current year Nk^t than 18 at flago-km.

Common significant factors for three studied TBE natural foci were relative air humidity of November of preceding year, relative humidity of April and June as well as temperature of June of current year that were revealed by means of decision tree construction. Unidirectional positive correlation between solar activity expressed in Wolf's numbers and TBE rate was shown for Novosibirsk and Irkutsk but for Gorno-Altaysk such correlation after 2001 was absent. Significant influence of vaccination on TBE prevalence was not found. As a result, more 40 logical regularities were obtained [1, 7], for example, the correlation between the decrease of questing ticks in the summer and the combination of air temperatures and humidity in the form of saturation deficit.

Conclusion

The methods of obtaining logical regularities based on analysis of multidimensional samples were proposed. The toolkit was used for the following applications: analysis of environmental factors influencing the incidence of tick-borne encephalitis in endemic regions of Russia. Suggested methods have some positive features. For example, obtained results are similar to natural language and reflect cause-effect relations of objects under investigation. The methods were tested on astrophysical, climatic data, and tick-borne encephalitis indices. They are especially useful in areas, where few data can be collected. Applications for suggested methods are possible in many areas, including, for instance, diagnostics, business, and quality control.

References

1. Analysis of Factors Influencing the Incidence of Tick-Borne Encephalitis, Using Logical-and-Probabilistic and Correlation-Regression Models / V.B. Berikov [et al.] // *Epidemiology and Immunization*. – 2011. – Vol. 6 (61), – P. 25–33. [In Russian]
2. Gunther, G. Tick-borne encephalopathies: epidemiology, diagnosis, treatment and prevention / G. Gunther, M. Haglund // *CNS Drugs*. – 2005. – Vol. 19 (12). – P. 1009–1032.
3. Gritsun, T.S. Tick-borne encephalitis / T.S. Gritsun, V.A. Lashkevich, E.A. Gould // *Antiviral Research*. – 2003. – Vol. 57 (1–2). – P. 129–146.
4. Influence of climatic factors on dynamics of questing Ixodes ricinus ticks in Slovenia / N. Knap [et al.] // *Vet Parasitol*. – 2009. – Vol. 164 (2–4), – P. 275–281.
5. Lbov, G.S. Logical decision functions and problems of statistical stability of decisions / G.S. Lbov, N.G. Startseva. – Novosibirsk : SB RAS Inst. of Mathematics, 1999. – 212 p. [In Russian]
6. Lbov, G.S. Stability of decision functions in problems of recognition and analysis of heterogeneous information / G.S. Lbov, V.B. Berikov. – Novosibirsk : SB RAS Inst. of Mathematics, 2005. – 217 p. [In Russian]
7. Study of the influence of natural factors on incidence of tick-borne encephalitis / G.S. Lbov [et al.] // *Bulletin of the NSU. Series: biology, clinical medicine*. – 2010. – Vol. 8 (2). – P. 31–37. [In Russian]
8. Mantke, O.D. A survey on cases of tick-borne encephalitis in European countries / O.D. Mantke, R. Schädler, M. Niedrig // *Euro Surveill*. – 2008. – Vol. 13 (17).
9. Singlair, A.R.E. Solar activity and mammal cycles in the northern hemisphere / A.R.E. Singlair, J.M. Gosline // *The American Naturalist*. – 1997. – Vol. 149 (4). – P. 776–784.
10. Ye, L. Time series shapelets: a novel technique that allows accurate, interpretable and fast classification / L. Ye, E. Keogh // *Journal Data Mining and Knowledge Discovery*. – 2011. – Vol. 22 (1). – P. 149–182.

ELIMINATION OF THE NON-POSITIVE DEFINITENESS OF MATRICES OF PAIRWISE COMPARISONS

D. Pshenichny, S. Dvoenko
Tula State University, Russia
e-mail: sergedv@yandex.ru, denispshenichny@yandex.ru

It is often necessary to analyze data given by the mutual comparisons between objects. These objects are naturally represented by the matrix of pairwise comparisons. The absolutely most of the data analysis methods are based on the immersion of objects in some metric space. This immersion is possible, if the matrix of pairwise comparisons is positively definite. If not, we have to correct this matrix to eliminate its non-positive definiteness. An approach for minimal corrections of the non-positively definite matrices of pairwise comparisons is demonstrated here.

Introduction and the problem statement

Let us have the experimental data presented by pairwise comparisons between elements of the limited set. In this case, the experimental results can be presented as a square matrix of pairwise comparisons. If this matrix is positively definite, then these objects can be immersed as vectors with common origin in some metric space. In this case, similarity between objects is equivalent to scalar product of two corresponding vectors and dissimilarity between objects is equivalent to distance between two vectors [1].

However, pairwise comparisons sometimes form a non-positively definite matrix. To process such data mathematically correct we need to eliminate its non-positive definiteness. This can be done, for example, by using discrete Karhunen-Loeve decomposition that allows removing of contribution of negative eigenvalues [2]. Nevertheless, this method isn't acceptable sometimes because we can't control the scale of corrections in data. It would be good to recognize which objects contribute the largest violation in metrics and correct only them. We also would like to get corrected positively definite matrix, which is maximally similar to the initial one.

Let a normalized matrix $S(n,n)$ of pairwise comparisons be given. The main diagonal of this matrix consists of units and all other elements are symmetrical relative to it with absolute values of them less than 1. Our goal is to develop a technique that gives a normalized positively definite matrix where its deviation after corrections from the matrix S is minimal.

1. Object ranging by their contribution in metric violations

Let the matrix $S(n,n)$ be given. It has diagonal of units, its other elements are symmetric with respect to main diagonal and their absolute value is less than 1. Hence, S is the matrix of some quadratic form.

To see whether it is positively definite, we can use the Sylvester's criterion: the matrix of a quadratic form is positively definite, if and only if its principal minors are positive [3].

According to the consequence from the Sylvester's inertia law [3], the number of negative eigenvalues ν in S equals to the number of sign changes in the sequence $S_0 = 1, S_1, \dots, S_n$. Here the notation S_k means k^{th} principal minor of S . One can prove that simultaneous permutation of two rows and two corresponding columns in S does not change its eigenvalues. This transposition means transposition of the corresponding two elements of the set.

As it was shown in [4], values of the principal minors of S are descending from 1 and once some minor S_k can become negative.

The idea of optimal object ranging consists in finding permutation that minimizes the additional metric violations after correction of a current negative principal minor S_k . If violating objects will be concentrated at the end of the sequence of the principal minors, then the additional metric violations will be minimized. As a result, we can get the optimally corrected positively definite matrix.

In other cases for non-optimal sequence of principal minors it appears that most of all minors from S_k to S_n should be corrected. Therefore, it is better to have the first negative minor as far as possible from the beginning of their sequence.

Inertia law says that k can't be more than $n - \nu$, hence, we want to have k as close as possible to $n - \nu$.

The idea of shifting k to $n - \nu$ consists in following. We want to get such the ranging of the set elements, so every following set element causes minimal descending of the principal minor value.

This ranging can be built in following way. Let us have n set elements and the corresponding matrix $S(n, n)$ of pairwise comparisons with ν negative eigenvalues. Then we remove such set element so the matrix $S(n-1, n-1)$ has $\nu - 1$ negative eigenvalues and its determinant has the maximal absolute value. We assign n^{th} position in our ranging to this set element.

Because of changing of evenness of the negative eigenvalues, the determinant changes its sign. If it is not possible to find such element that changes the determinant sign, then we remove the element that brings the maximal absolute value of the determinant and this element obtains n^{th} position in the ranging. Then we repeat this procedure for the matrix $S(n-1, n-1)$.

Each iteration decrements size of the matrix S and increments the number of set elements that obtained their position in the ranging. Hence, when we get the matrix $S(1, 1)$ all the set elements will be ranged.

If we choose the set elements in some other order, the determinant of the corresponding matrix of pairwise comparisons will not be changed. Therefore, our ranging is invariant under the order of the set elements.

2. Row and column correction

According to the idea of optimal permutation, we assume that when we correct in the matrix its k^{th} row and k^{th} column, its $(k-1)^{\text{th}}$ principal minor is positively definite and the value of k^{th} principal minor is negative.

Let us evaluate the value of k^{th} principal minor using Laplace's determinant decomposition. Decomposing it by k^{th} row and k^{th} column subject to matrix symmetry we obtain the next equation:

$$\det S_k = \det S_{k-1} - \sum_{i=1}^{k-1} \sum_{j=1}^{k-1} (-1)^{i+j} s_{ki} s_{jk} (S_{k-1})_i^j.$$

Here the notation $(S_{k-1})_i^j$ denotes the $(k-1)^{\text{th}}$ principal minor where i^{th} row and j^{th} column were removed. In this case, the notation $(S_{k-1})_i^j$ is the value of this minor too. Now let us denote k^{th} row and k^{th} column elements as some unknown values which we should calculate: $x_i = s_{ki} = s_{ik}$, $i = 1, \dots, k-1$. Replacing s_{ki} and s_{jk} in the previous equation and applying equality for elements of an inversed matrix

$(s_{ij}^{-1})_{k-1} = \frac{(-1)^{i+j} (S_{k-1})_i^j}{\det S_{k-1}}$ we obtain the following:

$$\det S_k = \det S_{k-1} - \det S_{k-1} \sum_{i=1}^{k-1} \sum_{j=1}^{k-1} x_i x_j (s_{ij}^{-1})_{k-1}.$$

Let us note that the second term in the right part of this equation is the quadratic form of the inversed $(k-1)^{\text{th}}$ principal minor. We can make following conclusions from this equation:

1. Because of positive definiteness of S_{k-1} , the positive definiteness of $(S_{k-1})^{-1}$ follows. Quadratic form of a positively definite matrix is greater or equal than zero and it equals zero, iff for all $x_i = 0$. Then $\det S_k \leq \det S_{k-1}$, and $\det S_k = \det S_{k-1}$, iff for all $x_i = 0$;

2. If $\sum_{i=1}^{k-1} \sum_{j=1}^{k-1} x_i x_j (s_{ij}^{-1})_{k-1} > 1$, then $\det S_k < 0$.

Let us correct k^{th} row and k^{th} column elements so that the value of the determinant of the corrected matrix equals to some number within interval $0 < C \leq \det S_{k-1}$.

We need to solve the equation $\sum_{i=1}^{k-1} \sum_{j=1}^{k-1} x_i x_j (s_{ij}^{-1})_{k-1} = 1 - \frac{C}{\det S_{n-1}}$ for $x_i, i=1, \dots, n-1$. The right

part of this equation is a constant. Let us denote it as $\phi = 1 - \frac{C}{\det S_{n-1}}$.

This equation has the following geometric interpretation. This is a hyper ellipsoid which is an equipotent line with value ϕ at a hyper paraboloid $\sum_{i=1}^{k-1} \sum_{j=1}^{k-1} x_i x_j (s_{ij}^{-1})_{k-1} = z$. This hyper paraboloid is the geometric locus of all points \mathbf{x} , where z is the value of determinant of the corresponding matrix. Let us denote the initial point as a vector $\mathbf{x}_0 = (s_{1k}, \dots, s_{(k-1)k})$, which also belongs to this hyper paraboloid.

To solve this problem optimally, we need to find the closest point on the hyper ellipsoid to the initial point. Therefore, we should solve an optimization problem:

$$\begin{cases} f(\mathbf{x}) = \sum_{i=1}^{k-1} (x_{0i} - x_i)^2 \rightarrow \min \\ g(\mathbf{x}) = \sum_{i=1}^{k-1} \sum_{j=1}^{k-1} x_i x_j (s_{ij}^{-1})_{k-1} - \phi = 0. \end{cases}$$

Here $f(\mathbf{x})$ is a target function. It determines squared Euclidean distance between initial and target points to be minimized. Constraint $g(\mathbf{x})$ is a target determinant value.

Let us solve this problem using the method of Lagrange multipliers. First, we have to construct the Lagrange function: $F(\mathbf{x}, \lambda) = \sum_{i=1}^{k-1} (x_{0i} - x_i)^2 + \lambda \left(\sum_{i=1}^{k-1} \sum_{j=1}^{k-1} x_i x_j (s_{ij}^{-1})_{k-1} - \phi \right)$. Then we differentiate this

function of \mathbf{x} : $\frac{\partial F}{\partial x_t} = -2(x_{0t} - x_t) + 2\lambda \sum_{i=1}^{k-1} x_i (s_{it}^{-1})_{k-1}$, $t=1, \dots, k-1$. Therefore, we obtain the set of k algebraic equations over k variables:

$$\begin{cases} x_{0t} - x_t = \lambda \sum_{i=1}^{k-1} x_i (s_{it}^{-1})_{k-1}, \quad t=1, \dots, k-1, \\ \sum_{i=1}^{k-1} \sum_{j=1}^{k-1} x_i x_j (s_{ij}^{-1})_{k-1} = \phi. \end{cases}$$

This system can have zero, one or more solutions in real numbers. If it has solutions, then the target point is among them. If not, it means the target determinant value is larger than maximal determinant value.

We are going to solve this set of equations using standard Newton iterations method which needs the initial value (we use \mathbf{x}_0 for \mathbf{x} and 0 for λ). It converges to the nearest solution. According to a meaning of a solution, we expect to obtain the nearest point to the initial one. Most probably, solution method converges to this value.

3. Example of experimental results

Result of the optimal ranging of objects is presented in [4]. Here we show the result of the optimal row and column correction. Let us test this method for a matrix

$$S(3,3) = \begin{pmatrix} 1 & 0.5 & 0.5 \\ 0.5 & 1 & -0.9 \\ 0.5 & -0.9 & 1 \end{pmatrix}.$$

Its determinant is -0.76 , so it is non-positively definite matrix. Its second principal minor is positively definite matrix with determinant value $\det S_2 = 0.75$. We'd like to correct the third row and column so that the determinant of the corrected matrix becomes equal to 0.1 . First, we need to calculate the inversed matrix $S(2, 2)^{-1} = \begin{pmatrix} 4/3 & -2/3 \\ -2/3 & 4/3 \end{pmatrix}$. Then we obtain the set of equations:

$$\begin{cases} \lambda(\frac{4}{3}x_1 - \frac{2}{3}x_2) = 0.5 - x_1 \\ \lambda(-\frac{2}{3}x_1 + \frac{4}{3}x_2) = -0.9 - x_2 \\ \frac{4}{3}x_1^2 - \frac{4}{3}x_1x_2 + \frac{4}{3}x_2^2 = 1 - \frac{0.1}{0.75}. \end{cases}$$

Solving this set of equations we obtain the root: $x_1 = 0.285487, x_2 = -0.624637$. Replacing matrix elements $s_{31} = s_{13}$ with x_1 and $s_{32} = s_{23}$ with x_2 we obtain corrected matrix which determinant equals to 0.1 . Euclidean distance between the initial and corrected matrices equals to 0.349057 .

Conclusion

Here we developed a new method of minimal corrections of the non-positively definite matrices of pairwise comparisons between some objects. This method allows objects ranging by their contribution in metric violations and correcting only of the most violating objects using the optimal row and column correction method.

This work is supported by the RFBR Grant 13-07-00010.

References

1. Young, G. Discussion of a set of points in terms of their mutual distances / G. Young, A.S. Householder // *Psychometrika*. – 1938. – Vol. 3 (1). – P. 19–22.
2. Pekalska, E. The Dissimilarity Representation for Pattern Recognition. Foundations and Applications / E. Pekalska, R.P.W. Duin. – World Scientific : Singapore, 2005. – 607 p.
3. Gantmacher, F.R. The Theory of Matrices / F.R. Gantmacher. – AMS Chelsea Publishing, 2000. – 660 p.
4. Dvoenko, S.D. Localization of the negative eigenvalues for a non-positively definite matrix of pairwise comparisons / S.D. Dvoenko, D.O. Pshenichny // *Proc. of the 11th Intern. Conference on Pattern Recognition and Image Analysis (PRIA-11-2013)*. – Samara, 2013. – P. 87–89.

TARGETS VISUALIZATION TOOL TO EXPLORE TESTING RESULTS

M. Semenov, D. Semenov
Tomsk Polytechnic University, Russia
e-mail: sme@tpu.ru

The article describes Target visualization tool. The Target is an innovative pie chart, in the interface of which the target metaphor is used. The tool is intended for exploration of knowledge testing results and competences monitoring in framework of Open Social Student Model. Using this tool, students can compare each other's results, which motivates more productive learning.

Introduction

At present the use of social learning technologies in education has become important in modern e-learning. However, the learning evaluation and assessment turns out to be a challenge. For evaluation and assessment, Hsiao I.-H. et al [1] propose to use Open Social Student Model (OSSM). This method gives students an opportunity to observe their current results making it possible to plan the trajectory of their learning. Bull S. et al [2] have demonstrated that the use of OSSM motivates students to communicate and work in teams as well as stimulates their anticipatory learning.

Speaking about researchers and tutors, it is important to realize for them what impact the cognitive tools of visualization may have on students' academic achievements. Vassileva J. [3] have studied the role of visualization among the internet communities and proved that visual presentation of personal results vs the internet community results increases the interaction among students, facilitates competitiveness and influences the team confidence positively.

To visualize results of testing, learning and monitoring, the authors introduce various tools such as Parallel IntrospectiveViews [1], CourseVis [4], Progressor [5], 3-simplex [6], eMUSE [7]. With the help of these tools, students can visualize their achievements and compare them with the achievements of other team members.

The purpose of our research is to develop and pilot the Target tool. To achieve the goal the following should be done: a) to analyze existing visualization tools; b) to describe main characteristics of the suggested tool; c) to determine the procedure of functional verification and tool testing

All of these form the area of the research. In the following chapter, the brief overview of different approaches of tool visualization development with regard to Open Social Student Model and social learning is given as well as the tool potential is described.

Definitions, the interface and properties of the Target tool are described in Section 2. In Section 3, preliminary steps to assess Target tool properties and to evaluate the degree of their influence on students are introduced. In the conclusion, the results are summarized and further research is discussed.

1. Related work

Information processing is a trend in the field of Computer Science. One of the primary missions of this trend is the development of methods for submission of great amounts of abstract data. The data should be presented visually to be interpreted by users. R. Spence [8] points out three main goals of visualization: exploration (search for relations, revealing regularities, decision-making, tendencies and anomalies); confirmation (verification of hypotheses), as well as presentation (submission of information). Hsiao I.-H. et al [1, 5], Falakmasir M.H. et al [9] notice that information visualization is an effective method which is actively used in the modern e-learning. Visualization instruments can be useful for both lecturers and students. However, the majority of investigations in that field are directed to the development of visualization instruments for lecturers and administration, while the minor part of projects is oriented on the investigation of visualization instruments for students [10–13]. On the other hand, intelligent training and testing systems [14] are equipped with components which provide students with knowledge (information) based on monitoring the students' behavior within the system. In traditional models of study students are unable to use this knowledge. Research [10-13] has shown the advantage of the Open Student Model, namely the model in which the student's learning status (knowledge, difficulties, misconceptions, etc.) is displayed visually. Moreover, the Open Student Model can be expanded to the Open Social Student Model [1]. In this model student has access to other students' learning statuses, which is especially relevant for social learning technologies. In this type of model the student can

compare his or her own results with the group's average results as well as with the best results of their internet-community.

Mitrovic A. [15] specify two main trends in implementing the OSSM. The first one focuses on imaging for the support of self-reflection and planning. The second trend encourages students' teamwork and cooperation. Chen et al. [16] investigated OSSM in order to motivate students to improve their academic achievements. Both individual and group open students models were studied; they both demonstrated an increase in reflection and helpful interactions among team mates.

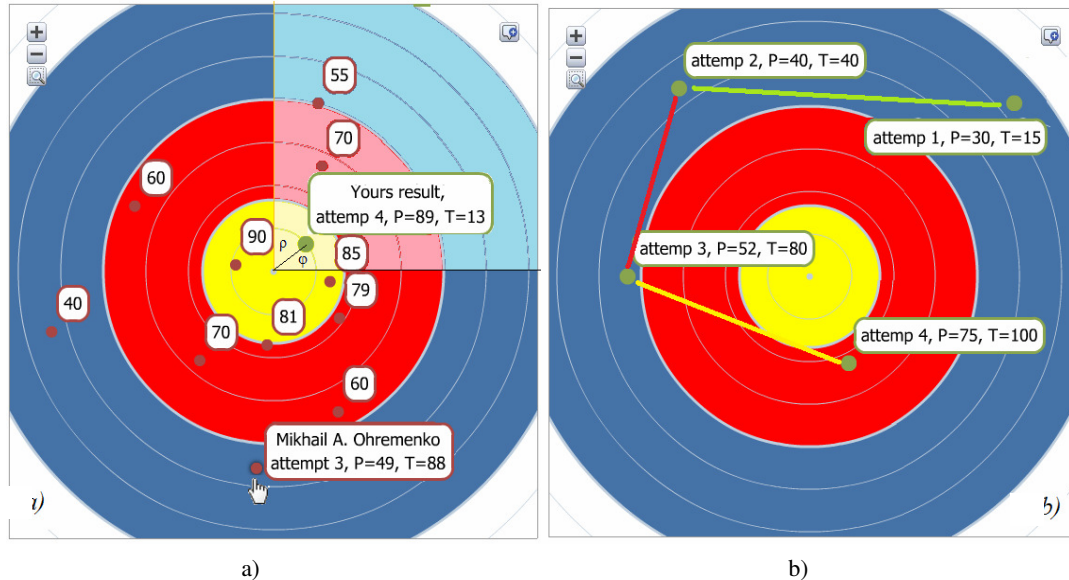


Fig. 1. Target visualization tool: a) display mode, b) trajectory mode

Bull S. [17] note that using the OSSM makes it possible to maintain students' metacognitive activities such as reflection, planning and self-evaluation, moreover, it improves students' awareness, motivates learning and reflection of knowledge. To achieve the required mastery level, the student can repeat the test session with various tasks offered. Strong students use various methods of enhancing their cognitive activity (e.g. advance independent work), while weaker students follow them [5].

Tools for delivering results of testing, learning, monitoring vary from simple (e.g. friends' walls on Facebook and charts [7]) to complex concept maps [5] and Bayesian network [18].

Publication [1] presents the Parallel Introspective Views tool. Using this instrument allowed increasing the average amount of time spent on self-assessment quizzes by 20%. Moreover, the average success rate (the number of the correct answers divided by the number of attempts) also increased by 9%. After such positive results, the authors are planning to integrate the proposed visualization tool in popular online learning environments, for example, Moodle. CourseVis [4] is system providing visualization for multiple groups of users to teachers and students. It helps tutors to identify problems in e-learning early on. A cognitive graphics tool, based on a 3-simplex [6] is appropriate to use both for decision-making and decision justification as well as to decision some of the following problems in intelligent training and testing systems. Progressor [5] is a visual interface for OSSM. It provides students the interface on their progress and allows comparing it to the classmates progress. eMUSE [7] is a platform which offers the tracking and visualization support needed for social e-learning. The platform collects students' actions from social media tools (e.g., blog, wiki, YouTube, Twitter).

2. Target – An Innovative Pie Chart

For the first time the Target visualization tool was presented in [19], this paper is the development of the research. Proposed Target tool (fig. 1) was designed to visualize test results and can be considered as an instrument for the OSSM. Target combines the basic trends of using the OSSM. On the one hand, Target displays the process of accumulation of knowledge, on the other hand, it enables comparing one's personal trajectory with achievements of the internet community.

Let us introduce the necessary notation. Let t_i be the planned completion time for the i -th task, $i=1, 2, \dots, n$, n is the number of tasks in one testing session. Then the total time for the testing session can be calculated using the following formula:

$$T_{\max} = \sum_{i=1}^n t_i . \quad (1)$$

Let P_{\max} be the maximum number of points that can be scored in one session. Generally, $90 \leq T_{\max} \leq 120$ minutes, $P_{\max} \leq 100$ points. The time required to complete each task is estimated by an expert. Commonly $t_i \neq t_j \forall i \neq j$ because test items have varying difficulty. For future discussions it is sufficient to use two indicators: the number of points (P) and time taken to complete the task (T). These indicators are available after the completion of the test session. Scores obtained in the past testing sessions are not summed. Note that if necessary, P_{\max} can be converted into relative units and the result can be measured in percentage units.

We suggest using the metaphor of a target and display the obtained results on the pie chart as markers (fig. 1). The default colour of a marker is red. The centre of the chart has coordinates (0, 0) in the polar system, the chart radius is P_{\max} . The coordinates of each marker depends on the number of points scored (polar radius $\rho > 0$) and the task completion time (polar angle $0 \leq \varphi \leq 2\pi$). Polar radius ρ and polar angle are denoted in fig. 1. The marker's radial coordinate ρ corresponds to the distance from the point of origin to the marker, and the polar angle φ equals the angle needed to turn the polar axis counterclockwise to catch this marker:




$$\begin{cases} \rho = P_{\max} - P, \\ \varphi = \frac{360 \cdot T}{T_{\max}}, \end{cases} \quad (2)$$

where P is the number of points scored in the given testing session, T is the time taken for testing session measured in minutes. Thus, the better the test result in the number of points scored, the closer the marker to the centre of the chart. As T grows, the polar angle φ also increases.

Fig. 1 shows that the chart consists of concentric circles. Each circle corresponds to a certain range of scores. If the score is from 80 to 100, the circle is yellow, if the score is between 55 and 79, it is red, etc. In addition to that, four sectors can be displayed on the chart, each sector sized $\pi/2$. The first sector $0 \leq \varphi < \pi/2$ contains results obtained within 30 minutes, the second sector $\pi/2 \leq \varphi < \pi$ contains results obtained within the range from 31 to 60 minutes, etc. The number of sectors on the chart is subject to variation. In fig. 1a the cursor is pointed to the marker with coordinates (49, 265). The popup window states the testing result: student Mikhail A. Ohremenko has scored 49 points in 88 minutes in the 3rd attempt. The current user's results are displayed with a green marker and a popup window showing the attempt number, time taken in minutes and the number of points scored.

Let us list the main features of Target tool.

Interactivity. Students can interact with the content of the Target as well as with other participants.

Interactivity is available in different modes of result presentation: display mode (fig. 1, a) and trajectory mode (fig. 1, b). The display mode is employed by default. At the user's request the results of other participants can be hidden or displayed on the chart. For example, it is possible to display on the screen both information about the results and contact data of another user. For that, it is necessary to press the button  "Info". The Target tool can be scaled using the buttons  "Zoom In" and  "Zoom Out", which allow detailing information on a certain segment of the chart.

Sequence. The Target chart in the trajectory mode presents the sequence of test results (fig. 1, b). Using this feature the user can analyze the process of knowledge growth and plan the future trajectory of learning and competence formation. In fig. 1, b the line connects four markers the coordinates of which are displayed in the popup windows. The different fragments of this line are differently coloured. The students receive the feedback using the traffic lights metaphor. The green colour shows there was one day or more between the first and second attempts, the yellow colour shows the period between attempts within one day, the red colour means the attempt was taken immediately after the preceding one. fig. 1, b shows that four testing sessions have occurred, each one resulting in the score higher than the previous one.

Identity. The student can establish the identity between the personal results and other participants' ones. This feature of the Target is the starting point for interaction inside the internet community.

Comparison. Students can compare each other's results, which motivates more productive learning. Another circle can be added to the chart, which shows the average score for all the participants. This

feature enhances interaction between the students, encourages competition and positively influences the microclimate of trust inside the group.

3. Evaluation & Pre-Study

In order to validate the claimed properties of the Target tool and to determine the degree of their influence on students' metacognitive activities (reflection, planning, self-assessment, motivation, knowledge reflection), the research has been initiated in the spring term 2014.

At present, the Target tool is used by the interactive informedia "Estimation of results and competences", Tomsk Polytechnic University (<http://exam.tpu.ru>). The Target tool has been realized using the multimedia platform Adobe Shockwave. The Target tool is designed to visualize the testing results of students' knowledge. In the interactive informedia, four subjects are introduced: Russian language, Mathematics, Physics and Chemistry. The tests bank contains more than 500 tasks for each subject. It allows to form a representative test sample for every testing session.

All registered users have got an access to their testing results through the Target interface. 43 out of 126 registered users actively exploit the available tool. The activity rate of the tool use is recorded. At every access, a user's name and timing are stored as well as the tool functions employed. In the future, students' questioning is going to be held for the purpose to get subjective opinion on the tool [1].

This work was financially supported by Russian Foundation of Basic Research (projects № 13-07-98037-r_sibir_a) and MSE Program "Nauka" (contract № 1.604.2011).

Conclusion

This article presents a review of trends in the development of innovative visualisation tools within the Open Social Student Model and social education.

The article describes the innovative Target visualisation tool. Its interface employs the target metaphor. Higher results are displayed as markers closer to the centre of the chart. The tool makes it possible to analyse the process of knowledge accumulation and plan the future personal strategy of studying and competence development. The tool interface is based on the Open Social Student Model. Two modes of result presentation are available: the display mode and the trajectory mode. Students receive feedback with the help of the traffic lights metaphor.

Using this tool, students can compare each other's results, which motivates more productive learning. The Target tool encourages students' metacognitive activities (reflection, planning and self-evaluation, reflection of knowledge) and increases their level of awareness. It is suggested that the Target tool should be used to present the results of study, monitoring and testing in e-learning and also to patterns recognition, to revealing different kinds of regularities, decision-making. The Target tool can be used to adaptive learning and as one of human identification approaches.

Our current results are encouraging and suggest new challenges for the future work. Further research is concerned with analysing the obtained data on using the Target tool and increasing the range of its functional features. In particular, it is planned to introduce the automatic calculation of proximity measures within the classmate or the internet community. In order to facilitate collaboration among students, we plan to add a feature for sending messages from the Target interface.

References

1. Open Social Student Modeling: Visualizing Student Models with Parallel Introspective Views / I.-H. Hsiao [et al.] // Proc. of the 19th Intern. Conference on User Modeling, Adaptation, and Personalization (UMAP-2011). – Girona, Spain, 2011. – Vol. 6787. – P. 171–182.
2. Bull, S. Group Interaction Prompted by a Simple Assessed Open Learner Model that can be Optionally Released to Peers / S. Bull, M. Britland // Workshop on Personalization in E-learning Environments at Individual and Group Level at the 11th Intern. Conference on User Modeling (UM-2007). – Corfu, Greece, 2007.
3. Vassileva, J. Evolving a Social Visualization Design Aimed at Increasing Participation in a Class-Based Online Community / J. Vassileva, L. Sun // International Journal of Cooperative Information Systems. – 2008. – Vol. 17 (4). – P. 443–466.
4. Mazza, R. CourseVis: Externalising Student Information to Facilitate Instructors in Distance Learning / R. Mazza, V. Dimitrova // Proc. of the Intern. Conference in Artificial Intelligence in Education (AIED-2003). – Sydney, Australia, 2003. – P. 279–286.

5. Comparative Social Visualization for Personalized E-learning / I.-H. Hsiao [et al.] // Proc. of the 11th Intern. Working Conference on Advanced Visual Interfaces. – Naples, Italy, 2012. – P.303–307.
6. Yankovskaya, A.E. Application of Cognitive Graphics Tools in Intelligent Systems / A.E. Yankovskaya, A.V. Yamshanov, N.M. Krivdyuk // International Journal of Engineering and Innovative Technology. – 2014. – Vol. 3 (7). – P. 58–65.
7. Popescu, E. Instructor Support for Monitoring and Visualizing Students' Activity in a Social Learning Environment Computers and Information Technology Department / E. Popescu, D. Cioiu // Proc. of the 12th IEEE Intern. Conference on Advanced Learning Technologies (ICALT). – Rome, Italy, 2012. – P. 371–373.
8. Spence, R. Information Visualisation / R. Spence. – Addison-Wesley, 2001. – 368 p.
9. The Impact of Social Performance Visualization on Students / M.H. Falakmasir [et al.] // Proc. of the 12th IEEE Intern. Conference on Advanced Learning Technologies (ICALT). – Rome, Italy, 2012. – P. 565–569.
10. Repairing Disengagement with Non Invasive Interventions / I. Arroyo [et al.] // Proc. of the Intern. Conference on Artificial Intelligence in Education. – Marina del Rey, CA, 2007. – P.195–202.
11. Zhang, H Moodog: Tracking Student Activity in Online Course Management System / H. Zhang, K. Almeroth // Journal of Interactive Learning Research. – 2010. – Vol. 21 (3). – P. 407–429.
12. Long, Y. Students' Understanding of Their Student Model / Y. Long, V. Aleven // International Journal of Artificial Intelligence in Education. – 2011. – Vol. 67 (38). – P. 179–186.
13. Arnold, K. Purdue Signals and Learning Analytics / K. Arnold [Electronic resource] // Learning & Knowledge Analytics. – Mode of access : <http://www.learninganalytics.net/?p=28>. – Date of access : 17.04.2014.
14. Building Intelligent Interactive Tutors: Student-centered strategies for revolutionizing e-learning. – 2009. – USA : Morgan Kaufmann Publishers, Elsevier, 2009. – 467 p.
15. Mitrovic, A. Evaluating the Effect of Open Student Models on Self-Assessment / A. Mitrovic, B. Martin // International Journal of Artificial Intelligence in Education. – 2007. – Vol. 17(2). – P. 121–144.
16. Active Open Learner Models as Animal Companions: Motivating Children to Learn through Interaction with My-Pet and Our-Pet / Z.H. Chen [et al.] // International Journal of Artificial Intelligence in Education. – 2007. – Vol. 17 (3). – P. 217–226.
16. Bull, S. Student Models that Invite the Learner / S. Bull, J. Kay // International Journal of Artificial Intelligence in Education. – 2007. – Vol. 17 (2). – P. 89–120.
17. Zapata-Rivera, J.D. Visualizing and inspecting Bayesian belief models/ J.D. Zapata-Rivera, J.E. Greer // International Journal of Artificial Intelligence in Education. – 2004. – Vol. 14. – P. 1–37.
18. Yankovskaya, A. Cognitive Tool for the Representation of Test Results of the Blended Education / A. Yankovskaya, M. Semenov, D. Semenov // Collection of scientific papers Interactive System: Problem of Human-Computer Interaction. – Ulyanovsk, Russia, 2013. – P. 305–312.

MUTATION ANALYSIS OF GENETIC SEQUENCES AND BUILDING MUTATIONAL DATABASE IN CONTEXT OF TUBERCULOSIS TREATMENT

R.S. Sergeev, A.V. Tuzikov, I.S. Kavaliou
United Institute of Informatics Problems of the NAS of Belarus, Minsk
e-mail: roma.sergeev@gmail.com

Emergence of drug resistance has been recognized as a global threat since the era of chemotherapy began. This problem is extensively discussed in the context of tuberculosis treatment. Alterations in pathogen genomes are among the main mechanisms by which microorganisms exhibit drug resistance. Analysis of the reported cases and discovery of new resistance-associated mutations may contribute greatly to the development of new drugs and effective therapy management. The proposed approach allows to uncover the co-occurring genetic changes and assess their contribution to resistance phenotypes.

Introduction

Tuberculosis (TB) is a major public health threat in Belarus. The recent escalation of the occurrence of the disease has been complicated due to the appearance and development of multi drug-resistant tuberculosis (MDR TB) or extensively drug-resistant tuberculosis (XDR TB), as well as HIV/TB co-infection, which requires long-term treatment. The ability of TB agent to resist treatment is strongly connected with variations and mutations in specific parts of the bacteria genome. Information about resistance mutations and their relationships to each other may become very valuable for choosing adequate treatment regimen and preventing therapy failure.

Here we introduce a framework for analyzing TB genome data generated in our own research. Our final efforts are directed to creation a knowledgebase that will take into account the fact that information is constantly refreshing since more and more results become available due to other research projects in this field.

1. The Problem

We perform a genome-wide association study (GWAS) that seeks to identify single nucleotide polymorphisms (SNPs) in *M. tuberculosis* genome and check them for association with emergence of drug-resistance. Due to a large number of genetic and environmental variables that interact with each other, identifying a correlation between a genetic change and a phenotypic trait is not hard-and-fast prediction. Therefore, we can only statistically estimate the impact of any single genetic variant, such as a SNP, which often makes only a small contribution to an overall effect.

There are several lines of drugs used for TB therapy. Five first-line are usually used in combinations to treat initially infected patients that do not confer drug resistance, and the second-line drugs are used to treat severe cases of MDR-TB [1].

Practically, we compare DNA sequences isolated from two groups of organisms: drug-resistant (cases) and drug-susceptible (control). To identify SNPs the genomes are mapped against reference sequence H37Rv. Patient records containing treatment lab findings are considered as a source of information to classify between susceptible and resistant phenotypes.

Besides information obtained in our own research there is a lot of material available in the international TB databases, which can be integrated into resulting knowledgebase to provide more precise expertise.

2. Algorithms and Methods

Data analysis procedure comprises several steps or levels so that the output of each level is the input of the next one (fig. 1). At the lower level, new original pathogen genome sequences are analyzed. The first three steps are aimed to reveal population structure, find site covariations and identify signals of recent positive selection in target genes under certain conditions (e.g. specific drug or treatment regimen).

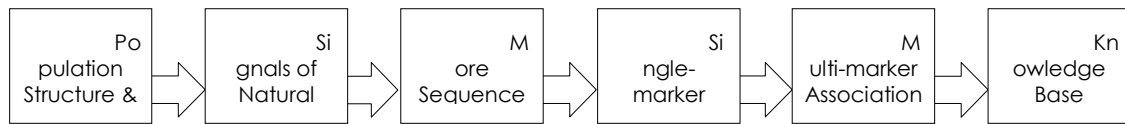


Fig. 1. Workflow of GWAS analysis

Next two levels of analysis uncover associations of genome variations with results of phenotype resistance tests to known drugs. And the last step is purposed to construct a probabilistic dependency network in order to structure the discovered associations and utilize information from external sources.

Population Structure and Genetic Diversity

The pipeline starts from identifying a population structure. We use phylogenic analysis by separate genes or whole genome to find out evolutionary relationships of the data to known clades of TB parasites. Principal component analysis (PCA) is another convenient technique to reveal population structure using genotyped parasite samples. The first two principal components are then taken to visualize results in conjunction with geographical origin of the sequenced samples.

Signals of Natural Selection

It is expected that parasite proteins which are involved in interactions with anti-TB drugs will be under positive or balancing selections to maintain diversity. It is important to recognize these regions because this information can support our hypothesis about associations that will be revealed further in statistical tests.

To measure genetic diversity we estimate nucleotide diversity statistic $SNP\pi$ as a measure of intrapopulation variation and Wright's population divergence statistic F_{ST} as a measure of the genetic distances between populations [2]. The evidence for positive selection can be examined by searching for regions with both high nucleotide diversity and high population divergence. In case of balancing selection the search is performed by regions with high nucleotide diversity and low population divergence [12].

Advanced Sequence Analysis

At this section we use more comprehensive approaches to reveal information from sequences like detecting correlated substitutions and haplotype phasing.

The goal of identifying correlated sites is to find out whether mutations in examined residues i and j are statistically associated. We use evolutionary and information theory approaches to reveal site covariations. The evolutionary approach is based on building evolutionary model, assessing it's parameters from the sequence alignment and counting for shared evolutionary events between examined sites [3, 4].

The sequenced data suggest that we have genotypes which can be easily derived from a list of SNPs. However, there are haplotypes that may become responsible for the observed variations in phenotypic traits of interest. Haplotypes are formed by gene alleles at adjacent loci that are in high linkage disequilibrium and inherited together. Haplotype phasing procedure assumes identifying haplotype blocks in the genome and estimating their frequencies based on genotypes observed in the alignment [5, 6].

Single-Marker Association Tests

Next levels of analysis are supposed to reveal associations of genome variations with results of phenotypic resistance tests to known drugs.

Most of single-marker tests rely on contingency tables approach which compare allele frequencies, or genotypes, in sets of cases and controls. However, standard statistical tests like Fisher's exact test or χ^2 -squared goodness-of-fit test generally show P-value inflation due to confounding effect from population structure [12]. Therefore we apply the permuted versions of these tests and more complex

techniques like Cochran-Mantel-Haenszel test [7] which can account for population structure through the specifying a cluster variable.

In case of multiple testing we always perform adjustment of P-value threshold. We use Benjamini-Hochberg approach [8] to control false discovery rate (FDR) that is a portion of errors made among alternative hypotheses at $q_{(k)} = \min_{t \geq p_{(k)}} FDR(t)$ [9], where $FDR(t) = E \left[\frac{F(t)}{S(t)} \right]$, $S(t)$ is the number of $p_{(i)} \leq t$, and $F(t)$ is the number of true null hypotheses among the accepted alternatives if the level of significance was set at $\alpha = p_{(k)}$.

Multi-Marker Association Tests

There is a discussion in literature about efficiency of single-marker tests in comparison to multi-marker tests. Some computations based on simulation studies suggest that single-marker tests provide at least as much power as multi-marker haplotype tests [10]. However, analysis of other GWAS researches shows that multi-marker tests can highlight potential associations that single-SNP techniques do not [11,12].

Though regression approach that relates probabilistically assigned predictors with trait values is not entirely new in bioinformatics, there is an increasing interest in using linear mixed models. We use two multi-marker tests relying on regression models to perform association analysis: efficient mixed model association (EMMA) and a haplotype likelihood ratio (HLR) test.

EMMA test and its modifications [13,14] identify associations of quantitative traits with SNPs along the genome in individuals with complex population structure adjusting for confounding effects from phylogeny and site covariations. This test is practically applicable for small sample sets and has realizations that efficiently compute exact values of test statistics. Within EMMA test we consider the following model equation:

$$y = X\beta + Zu + \varepsilon, \quad (1)$$

where n is the number of examined TB strains, y is a vector of observed phenotypes ($y_i = 1$ for drug-resistant and $y_i = 0$ for drug-susceptible), X is an $n \times q$ matrix of fixed effects including group means, SNPs, and other confounding variables, β is a vector representing coefficients of the fixed effects, Z is $n \times n$ identity matrix mapping each observed phenotype to each of n strains. Vector u denotes the random effect of the mixed model with $Var(u) = \sigma_g^2 K$, where K is the $n \times n$ similarity-based kinship matrix inferred from genotypes.

The HLR test is designed to detect association of a single haplotype within a phenotypic trait and reported to be particularly powerful when the analyzed haplotype experienced recent strong selection [12,15]. We assume that haplotypes and their frequencies have been resolved at sequence analysis step using sliding window with $t \leq 6$ SNPs, and the phenotypic trait $y \sim Bin(1, \pi)$ follows binomial distribution with probability π of drug-resistance phenotype. The phenotypic variable y is fit to the given predictor variable X of haplotype counts using logistic regression model:

$$\log \left(\frac{\pi}{1 - \pi} \right) = X\beta, \quad (2)$$

where β is a vector of haplotype effects. The null-hypothesis $H_0 : \beta_j = 0$ for the individual j -th haplotype is tested against alternative of haplotype significance and likelihood statistic $-2 \log(L_0/L_1)$ is calculated. The effect of population structure is addressed using population-specific permutations of the dependent variable.

Knowledge Base

At this level we are purposed to construct a probabilistic dependency network in order to structure associations discovered at the previous steps. As soon as the uncovered associations are based on probabilities, they can be represented as weighted arcs between variables. In this context, the variables in the model correspond to the presence or absence of amino acids in codons, received drugs, and treatment outcomes (fig. 2).

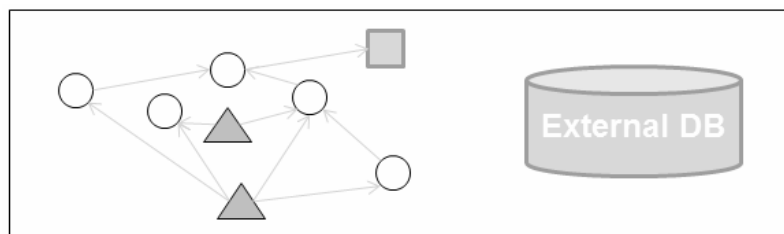


Fig. 2. Creating a dependency network using uncovered associations and external data

Information on drug resistance from other studies and public databases can be added at this level as supplementary associations in the network, which gives an advantage of taking into account all available data. Inference algorithms designed for Bayesian and Markov networks are used to retrieve information through queries to the networks composed from sets of observed and requested variables. The dependencies inside the network can be updated as soon as new data appear.

3. Results

TB genome sequencing is still in process, and we have tested elements of this approach on sample M.Tuberculosis data collected from public databases instead. Synthesized data were used as a source of assigned quantity traits (we assumed that phenotype dependency on genotype can be modeled using weight coefficients). As the result, in multi-marker association tests we observed rather good compliance with synthesized dependencies (fig. 3).

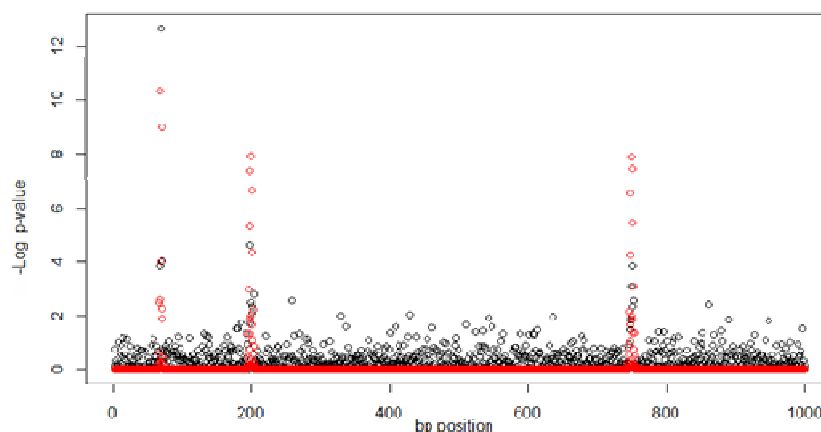


Fig. 3. Manhattan plot that demonstrates compliance of multi-marker EMMA test results (grey circles) with generated dependencies (red circles)

However, in more general case the choice of the methods for each step may significantly affect the outcome.

Conclusion

We can see that GWAS analysis pipeline involves many different steps and a range of statistical tests. There is third-party software and packages that perform some commonly used tests [16] and can be very helpful within the research. But they do not allow fully automate the process to get data for the knowledge base in the end. Therefore we still have to make efforts to integrate steps along the pipeline. Special attention should be given to setting up test parameters and interpretation of results.

Elements of this approach are used in current project performed in collaboration with NIAID of NIH through a CRDF BOB-31120-MK-13 project to establish the Belarus tuberculosis database (<http://tuberculosis.by>) and conduct comprehensive study of obtained MDR and XDR TB strains.

References

1. Ferguson, L.A. Multidrug-resistant and extensively drug-resistant tuberculosis: The new face of an old disease / L.A. Ferguson, J. Rhoads // *J Am Acad Nurse Pract.* – 2009. – Vol. 21 (11). – P. 603–609.
2. Nei, M. Mathematical Model for Studying Genetic Variation in Terms of Restriction Endonucleases / M. Nei, W.-H. Li // *PNAS.* – 1979. – Vol. 76 (10). – P. 5269–5273.
3. Coevolving protein residues: maximum likelihood identification and relationship to structure / D.D. Pollock [et al.] // *J. Mol. Biol.* – 1999. – Vol. 287 (1). – P. 187–198.
4. Sergeev, R.S. Algorithms for mutation analysis of HIV-1 subtype A primary protein sequences / R.S. Sergeev, A.V. Tuzikov, V.F. Eremin // *Informatics.* – 2011. – Vol. 3 (31). – P. 88–97.
5. Stephens, M. A new statistical method for haplotype reconstruction from population data / M. Stephens, N.J. Smith, P. Donnelly // *Am J Hum Genet.* – 2001. – Vol. 68. – P. 978–989.
6. Chiano, M.N. Fine genetic mapping using haplotype analysis and the missing data problem / M.N. Chiano, D.G. Clayton // *Am J Hum Genet.* – 1998. – Vol. 62. – P. 55–60.
7. Wittes, J. The Power of the Mantel-Haenszel Test / J. Wittes, S. Wallenstein // *Biometrics.* – 1993. – Vol. 49 (4). – P. 1077–1087.
8. Benjamini, Y. [Controlling the false discovery rate: a practical and powerful approach to multiple testing](#) / Y. Benjamini, Y. Hochberg // *Journal of the Royal Statistical Society, Series B.* – 1995. – Vol. 57 (1). – P. 289–300.
9. Storey, D.J. Statistical significance for genomewide studies / D.J. Storey, R. Tibshirani // *PNAS* 100. – 2003. – Vol. 16. – P. 9440–9445.
10. Kaplan, N. Issues concerning association studies for fine mapping a susceptibility gene for a complex disease / N. Kaplan, R. Morris // *Genet Epidemiol.* – 2001. – Vol. 20. – P. 432–457.
11. Genetic analysis of case/control data using estimated haplotype frequencies: Application to APOE locus variation and Alzheimer's disease / D. Fallin [et al.] // *Genome Res.* – 2001. – Vol. 11. – P. 143–151.
12. Identification and Functional Validation of the Novel Antimalarial Resistance Locus PF10_0355 in *Plasmodium falciparum* / D. Tyne [et al.] // *PLoS Genet.* – 2011. – Vol. 7 (4). – P. 1371–1383.
13. Efficient Control of Population Structure in Model Organism Association Mapping / H.M. Kang [et al.] // *Genetics March.* – 2008. – Vol. 178 (3). – P. 1709–1723.
14. Zhou, X. Genome-wide efficient mixed-model analysis for association studies / X. Zhou, M. Stephens // *Nature Genetics.* – 2012. – Vol. 44 (7). – P. 821–824.
15. Testing Association of Statistically Inferred Haplotypes with Discrete and Continuous Traits in Samples of Unrelated Individuals / D.V. Zaykin [et al.] // *Hum Hered.* – 2002. – Vol. 53. – P. 79–91.
16. [PLINK: a tool set for whole-genome association and population-based linkage analyses](#) / S. Purcell [et al.] // *Am J Hum Genet.* – 2007. – Vol. 81 (3). – P. 559–575.

PLANNING OF THE ENTERPRISE RESOURCES

G. Shangytbayeva

Kazakh National Technical University of a name of K.I. Satpayev, Almaty

e-mail: gul_janet@mail.ru

The article considers enterprise resource planning and its features, the program systems which is integrated within ERP systems.

Introduction

Enterprise Resource Planning (ERP) is information systems of planning of the enterprise resources, for a long time it became an ordinary field of activity for the enterprises of medium and large business. Implementation of the information networks is directed to increase efficiency of the business processes, that are represented by supply, sale and production. Implementation of this technology significantly influences to increase the productivity of the enterprise as a whole. But ERP is not only automation of business processes, it is also automation of such administrative functions, as planning, control and the account.

ERP – is uniform information system for management of work of all enterprise, with its variety of departments, with all their functions and characteristics. The main line of all ERP systems is the general database which supports a set of the functions used by various business by units. In an ideal ERP allows to the manager start the application and from one window receive information of any aspect of business, whether it be finance, human resource management or sale.

ERP has radical difference from the entire Microsoft Office program familiar to us which equally works at all computers. Functionality of ERP systems directly depends on accurate definition of tasks of the concrete enterprise and its control under these tasks. Full efficiency from use of this system is reached only if it is designed and adjusted correctly to the enterprise, so it will help to make further business more operated.

ERP-systems

Enterprise resource planning systems is used to integrate all business processes of the enterprise into the one uniform environment which will allow automating all necessary operations like control, the account, planning of production resources and the analysis of activity of the organization. Enterprise resource planning consists of a large number of program modules and also uniform database which stores in itself data on operations and enterprise business processes. These systems include planning, management and preparation of enterprise productions, management of warehouses work and a cargo transportation, the accounting of goods, purchases and deliveries, strategic business management, accounting, marketing and help service, possibility of management of an enterprise manpower, and also information technologies: Enterprise resource planning systems provide protection not only against malefactors from the outside, but also from internal, for example, they allow to prevent wastes or plunders at the enterprise.

Enterprise resource planning systems are constantly developing, but despite their obvious efficiency managers of some enterprises do not decide on their implementation. It occurs because implementation of such systems rather expensive, and not all companies are able to afford to allocate for it finance and to train the employees to work with this system, and also not all owners of the companies can trust at once in the uniform program comprising all information on a state of affairs on production, it seems to them unreal and unreliable. But the enterprises which implemented by itself ERP-system, understand all benefit of their use as ERP-system give to the enterprise advantages before competitors at the expense of competent maintaining expenses and achievement of the greatest efficiency of business processes.

Enterprise resource planning systems today are not only the powerful tool for management and business, but also serious advantage before competitors, thanks to that there is a coherence of work of various departments of the enterprise, reduction of administrative expenses and there is no need to transfer the necessary data from one program to another as all organization works with uniform system [1].

Distinctive feature of ERP-system from other systems is that it helps to aggregate data on activity of the enterprise and only on the basis of collected information the system will be able to analyze data.

Important feature of this system is that economic operations in system are registered the unique time, and we can analyze at once their influence on activity of the enterprise for the received reports.

Information system of planning of enterprise resources is quite developed system and its functions are in a constant stage of development and improvement, but nevertheless, it happens so that after ERP-system was implemented to the certain enterprise and all techniques of its implementation were involved correctly, the enterprise management still doesn't manage to receive complete information control over enterprise activity. And that the most interesting, anything essential doesn't occur, and even on the contrary, everything remains still. What is the matter here? The factors influencing to incorrect work of ERP-system can be much. It can be, for example, incorrect registration of primary documents, failures and violations in policy of sale, existence at the enterprise of excess stocks. Cases of are possible even that many enterprises after Enterprise resource planning introduction, refuse it for the reason that allegedly it inadequately and out of time reacts to the tasks set for it. But so kind of problems occur not only at our enterprises, there is an information, that is in the West less than 50% of cases are successful implementation of ERP-systems at their enterprises.

But why so large percentage of implementation of ERP-system is unsuccessful. If to carry out the analysis of unsuccessful introductions of Enterprise resource planning systems, the following clears up that violation of the principle of design of automatic control systems (ACS) appears one of the main factors of unsuccessful introductions. Among experts there is an opinion that projects of implementation of automated control systems do not yield positive results because at design of data of systems strategy of development of business isn't considered and too frequent reprogramming business processes is made.

ERP abbreviation is used by Enterprise-Resource Planning for designation of complex enterprise management systems. The mission of ERP systems consists in integration of all departments and company functions into uniform computer system which will be able to serve all specific needs of separate divisions.

The most difficult – to construct uniform system which will serve all inquiries of staff of finance department, and, at the same time, will please both to a human resources department, and a warehouse, and other divisions. Each of these departments usually has own computer system optimized under the features of work. ERP combines them everything within one integrated program which works with a uniform database, so, that all departments can communicate and communicate easier with each other. Such integrated approach promises to turn back very big return if the companies are able correctly to install system.

ERP replaces old separate computer systems on finance, human resource management, control over production, to logistics, a warehouse one unified system consisting of program modules which repeat functionality of old systems. The programs serving finance, production or warehouse are connected now together, and from one department it is possible to glance in information of another. Enterprise resource planning of the majority of suppliers are rather flexible and easily adjusted, they can be established modules, without getting at once all package. For example, many companies get at first only financial or HR modules, leaving on the future automation of other functions.

Enterprise resource planning automates the procedures forming business processes. For example, implementation of the order of the client: adoption of the order, its placement, shipment from a warehouse, delivery, drawing of the account, receiving payment. Enterprise resource planning “picks up” the order of the client and serves as some kind of road map on which various steps on a way of execution of the order are automated. When the representative of front office enters the order of the client into Enterprise resource planning, it has an access to all information, necessary to start the order for performance. For example, it right there gets access to a credit rating of the client and history of his orders from the financial module, learns about goods existence from the warehouse module and about graphics of shipment of goods from the logistics module.

The employees working in different divisions, see one information and can update it in the part. When one department finishes work on the order, the order is automatically readdressed in other department in the system. To learn where there was an order at any moment, it is necessary to enter only into system and to trace order passing. As all process is transparent now, orders of clients are carried out quicker and with smaller number of mistakes, than earlier. The same happens to other important processes, for example, creation of financial reports, salary charge, etc. [2].

In practice of management distribution was gained by the following program systems integrated within ERP systems into a single whole:

- SCM (Supply Chain Management) – management of supply chains;
- CRM (Customer Relationship Management) – management of relationship with clients;

- CRP (Capacity Requirements Planning) – planning of requirement for capacities;
- WMS (Warehouse Management System) – management of a warehouse – a control system providing automation and optimization of all processes of warehouse work of the profile enterprise. The system allows to operate warehouse logistics within various technological processes (reception and goods shipment, internal displacements) in real time. By means of automation of a warehouse high turnover of a warehouse is reached, the fast complete set of consignments of goods, shipment to their consumers is carried out.

Creation and Enterprise resource planning introduction in the company or at the enterprise initially assumes radical changes in business processes of this enterprise or the company. It is promoted by methodology of introduction of Enterprise resource planning, functionality of this Enterprise resource planning. There are class SAP R/3 systems for which the standard practice of introduction is pulling of business processes of the company to already available rich functionality of Enterprise resource planning. Owing to introduction of this system the enterprise incurs serious costs of management of the introduced project, and also because of changes in own business processes. Also there are such systems as MS Navision Attain and Axapta which it is considered to be systems of middle class. Data of system are one of the best systems in this class, at introduction of these systems the designer initially is guided by an adaptability to demanded processes of the enterprise. That is the main objective of introduction of these systems is focused on processes of the customer, and already then on functionality of system. Modern Enterprise resource planning possess the high-level and integrated environment of development, in this case creation of system which will meet fully all existing requirements of the company is possible [3].

Conclusion

Good Enterprise resource planning brings huge benefit to the enterprise and very quickly pays back the expenses connected with its introduction but only provided that introduction is finished and carried out correctly. And it is very difficult task. The numerous factors of the market arising as in the course of introduction of system, and its operation, demand the careful analysis and development of the measures preventing receiving unpredictable results.

After all difficulties and surprises wait continually, beginning from a real assessment of need of introduction of Enterprise resource planning at the enterprise, and finishing resistance of employees to work with already introduced system. To achieve desirable result, it is necessary nobility and remember already in a stage of planning of the project that in case of success the introduced control system will bring notable benefit.

In case of failure – will extremely negatively affect not only future prospects, but also at the current work of the enterprise.

References

1. Гуриев, В. Системы класс ERP для “чайников” / В. Гуриев, С. Питеркин. – Компьютера, 2001. – № 11.
2. Keller, D. Enterprise Resource Planning. The changing application model / D. Keller, L. Erik. – GartnerGroup, 1996.
3. CloudLinux [Electronic resource]. – Mode of access : <http://www.ccm.kz/article>. – Date of access : 17.04.2014.

MULTISCALE MOLECULAR SURFACE DESCRIPTORS

A. Shestov, M. Kumskov
Lomonosov Moscow State University, Russia
e-mail: shestov.msu@mail.ru

The paper is devoted to the methods of searching of quantitative structure-property relationships for predicting the activity of chemical compounds. The multiscale molecular surface description method is presented. It is based on our previous work about the molecular surface descriptors [1]. The new method improves descriptiveness and decreases the algorithm complexity.

Introduction

Quantitative structure-activity relationship problem [2] is a problem of finding dependencies between structure of chemical compounds and their activity. This problem is the application of pattern recognition in chemistry.

1. Basic definitions

Labeled molecular graph $G = \{E, V\}$ is a labeled graph which vertices are interpreted as atoms and edges are interpreted as valent bonds between them. Labels can be 3d coordinates, physical and chemical properties, etc. Let the learning set L_s contains N molecular graphs x_1, \dots, x_n , each molecular graph has activity value $y_i \in \{-1, 1\}$. Let us denote the set of molecular graphs with the labels of the same type as in learning set as Ch .

Let us name the set of functions $F : \{f : Ch \rightarrow \{-1, 1\}\}$ as classification functions set. The solution of the QSAR problem is the algorithm of classification functions construction:

$$a_{best} \in \{A : 2^{Ch} \rightarrow F\}, a_{best} = \max_{a \in A} q(a),$$

where $q(a) : A \rightarrow R$ is a quality functional.

The basic method for the QSAR problem solution is the following two-step method. Firstly each molecular graph x_i is represented as a vector v_i in R^n . Then for this set of vectors and activity values the problem is solved using general methods of pattern recognition. Vector v_i is called descriptors vector for molecular graph x_i .

2. Molecular surface descriptors

Let us describe the method of molecular surface descriptors construction. Its improvement will be discussed later.

Molecular surface [3] is a surface which is constructed by rolling a probe sphere around Vandervaals surface (fig. 1).

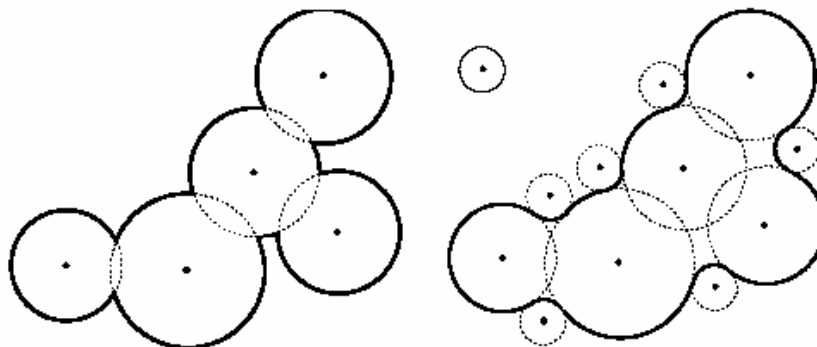


Fig. 9. Vandervaals (left) and molecular (right) surfaces

The so-called singular points are found on molecular surface. They are found as centers of regions of the same sign curvature (fig. 2).

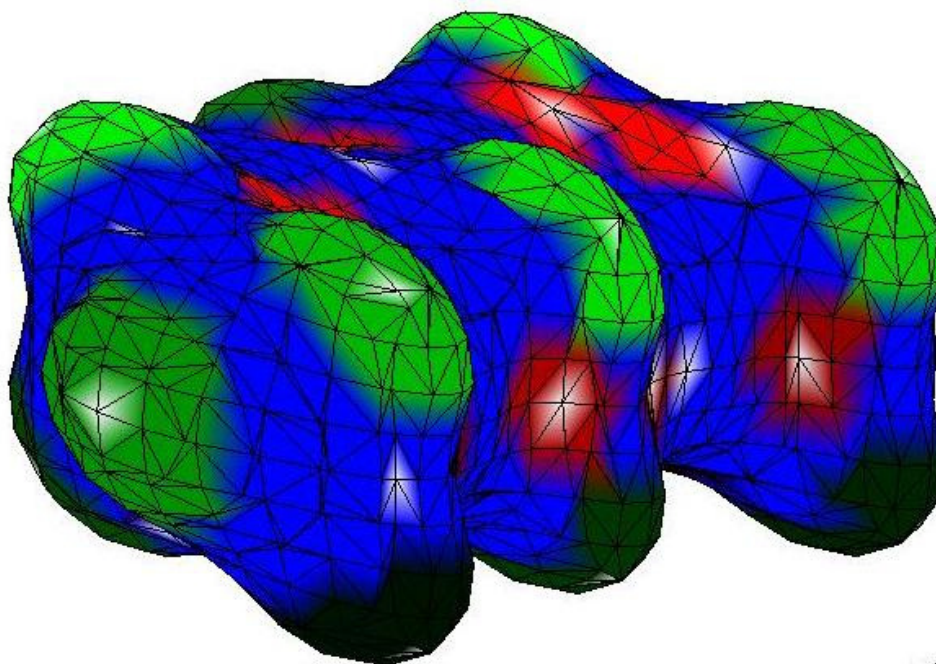


Fig. 10. A segmented molecular surface with marked singular points

Chemico-physical properties for each point are calculated. Pairs and triplets of singular points are formed. Points properties and distances between them are categorized. Based on this categorization, pairs and triplets of singular points are categorized. Then for each category the number of representing pairs and triplets is counted. The descriptors vector for molecular graph is formed out of these numbers for all categories.

This method has the following drawbacks:

- a molecular graph usually has about 200 singular points. So the number of triplets for each molecular graph is about 1300000. And calculation of singular points triplets and molecular descriptors takes a lot of time;
- triplets and pairs with large distances between singular points are poorly described. Categories of such triplets and pairs are not influenced by area between these singular points. So the structure of large fragments is not reflected in the categories of triplets and pairs.

3. Multiscale molecular surface descriptors

Now we present the improved method of molecular surface descriptors construction which don't has the mentioned drawbacks. We will call the constructed descriptors as multiscale molecular surface descriptors:

1. Firstly, we choose a set of scales $S = \{s_1, \dots, s_n\}$.
2. This set of scales defines a set of 3d Gaussian filters $G = \{g_1, \dots, g_n, \sigma(g_i) = f_s(s_i)\}$. We apply the set G to the coordinates of molecular surface and obtain a set of smoothed molecular surfaces $M = \{m_1, \dots, m_n\}$.
3. For each m_i a set of singular points $P_i = \{p_1, \dots, p_{j_i}\}$ is calculated. For large scales there will be less singular points and these points will correspond only to large geometrical structures.
4. Each singular point p_j on each surface m_i is described with a histogram of chemico-physical properties h_j . Histogram is calculated for a circle on the surface m_i . Circle is centered at the singular point and has radius $r_i = r(s_i)$.
5. Then pairs and triplets of singular points are formed. For each scale s_i there is a maximal distance between points d_i . So the number of points decreases a lot.

So this method doesn't have the drawbacks mentioned above:

1. Number of pairs and triplets at fine scales decreases a lot due to the restrictions of distance between points. Number of singular points at coarse scales decreases, because surface details are not

presented on filtered surface with large sigma.

2. Large fragments of molecular surface are described well, because both radius of histogram and distance between points depends on scale s_i . So all area between points influences on descriptor value.

Conclusion

A method for construction of multiscale molecular surface descriptors is presented. The method is an improvement of our previous work. The new algorithm has a lower complexity. The description of large structures on molecular surface was improved. The method can be applied in the QSAR and QSPR studies.

References

1. Shestov, A. The new molecular surface descriptors / A. Shestov, A. Koptsova, M. Kumskov // Proceedings of the 21th Intern. Conference of Computer Graphics and Vision Graphicon. – M. : MAKS Press, 2011. – P. P-17.
2. Varmuza, K. Pattern recognition in chemistry / K. Varmuza // Lecture Notes in Chemistry. – 1980. – Vol. 21. – P. 59.
3. Connolly, M.L. Solvent-accessible surfaces of proteins and nucleic-acids / M.L. Connolly // Science. – 1983. – Vol. 221. – P. 709–713.

THE COMPLEXITY OF INDUCTIVE RESOLUTION ALGORITHMS IN PATTERN RECOGNITION PROBLEMS

O. Shut

Belarusian State University, Minsk

e-mail: olgashut@tut.by

The paper considers three inductive resolution algorithms for solving pattern recognition problems in the case when all features of objects have a finite quantity of values. The computational complexity of all described algorithms is calculated. In order to reduce running time of algorithms, modifications, based on the idea of branch and bound method, are suggested for all three algorithms. Complexities of modified algorithms are estimated. The running time of inductive resolution algorithms is shown to decrease exponentially depending on the structure of information of the object being considered.

Introduction

The scheme of using resolution method for solution of supervised pattern recognition problems is suggested in [1]. An analog of resolution method, which processes objects in finite-dimensional space of feature values, is developed. Algorithm A_1 , which is called object resolution algorithm, and algorithm A_C , which combines A_1 with algorithm A_2 , are built. Algorithm A_2 is described in [2] and called inductive resolution algorithm in [3]. Because A_C also implements inductive derivation, A_C may be also classified as an inductive resolution algorithm. The main idea of these algorithms is search of objects, which can be derived with using of object resolution method from the given set or learning sample. Therefore it is necessary to reduce run times of these algorithms. For this purpose, application of the technique, introduced in [4], where closeness of feature values is estimated to reduce run time of A_2 due to using of branch and bound method, is suggested. This paper considers modifications of all listed algorithms, based on this method, estimates their complexity and shows that complexity of inductive resolution algorithms can be reduced.

1. Main concepts and problem definition

Concepts from [1-4] are used in this paper. Let us consider pattern recognition problem Z , where X denotes the set of objects, X_1, \dots, X_l denote classes and $X^0 \subset X$ denotes the learning sample. We also need some definitions: an object is a mapping $p: s_1 \times \dots \times s_n \rightarrow D^n$; $S = \{s_1, \dots, s_n\}$ is a set of features with fixed order; D is a finite set of their values; $k = |D|$; D_j^p is a set of values of the feature s_j of object p , where $D_j^p \subset D$, $D_j^p \neq \emptyset$, $j = \overline{1, n}$. For description of objects, code c , where the value of feature s_j of object p is represented as a k -dimensional vector, is used:

$$c(D_j^p) = (d_{j1}^p \quad \dots \quad d_{jk}^p), \quad d_{ji}^p = \begin{cases} 1, & (i-1) \in D_j^p \\ 0, & (i-1) \notin D_j^p \end{cases}.$$

Object p , which has n features, is represented as vector $c(p) = (c(D_1^p) \quad \dots \quad c(D_n^p))$. Let us also consider the operation of building of object resolvent $r = Or_h(p_1, p_2)$:

$$D_j^r = \begin{cases} D_j^{p_1} \cup D_j^{p_2}, & j = h \\ D_j^{p_1} \cap D_j^{p_2}, & j \neq h \end{cases} \quad \text{or in code } c: \quad c(D_j^r) = \begin{cases} c(D_j^{p_1}) \vee c(D_j^{p_2}), & j = h \\ c(D_j^{p_1}) \wedge c(D_j^{p_2}), & j \neq h \end{cases}.$$

Let us briefly describe algorithms A_1 , A_2 , and A_C . Detailed description of these algorithms can be found in [1]. All objects are assumed to being represented in code c . Algorithm A_1 considers sequentially all pairs of objects (p_1, p_2) from X^0 , where both objects are known to belong to the same class, and builds a new object $r = Or_h(p_1, p_2)$, which is also supposed to belong to the same class. New

objects are built until either belonging of required object x is determined or all pairs of objects have been considered. Operation of building of r is shown in [1] to correspond to the operation of building of resolvent in Boolean logic, hence A_1 is called object resolution algorithm.

To describe algorithm A_2 , let us introduce a function $\mu : (c(X))^2 \rightarrow [0,1]$:

$$\mu(x, y) = \max\{0, (\sum_{j \in S} (-1)^t a_{ij}) \cdot (\sum_{j \in S} a_{ij})^{-1}\}.$$

Algorithm A_2 uses function μ to calculate closeness of objects to each other. Parameters $\|a_{ij}\|$ are determined at the learning stage. A possible scheme of calculating of $\|a_{ij}\|$ is suggested in [2]. At the recognition stage $\mu(x, x^0)$ is calculated for all $x \in X$ and $x^0 \in X^0$, and its maximum is considered to determine the class of object x .

Algorithm A_C is a combination of algorithms A_1 and A_2 : first it executes algorithm A_1 for all $x \in X$, then it executes algorithm A_2 for all objects which belonging couldn't be determined by A_1 .

Let us estimate the computational complexity of these three algorithms. For this purpose let us consider three operations used in these algorithms: comparison of two objects, building of r and check of existing of object. It follows from their definitions and formulas that complexity of each of these three operations is $O(kn)$. Reasoning from these estimates, one can calculate that complexities of A_1 , A_2 , and A_C are $O(kQ^3n^2)$, $O(kq_0n^2)$ and $O(kQ^3n^2)$ correspondingly (tables 1-3), where $q_0 = |X^0|$, Q is the total quantity of built objects. Here X is assumed to be stored as a list.

Table 1

Complexity of algorithms A_1 and A_1^B

Step A_1	Step A_1^B	Complexity
Steps 1–3	Step 1	Steps 2–8 are executed n times
	Step 2	Steps 3–5 are executed l times
	Step 3	Steps 4–5 are executed q_i^2 times, where $q_i = Y_i $
Steps 4–5	Steps 4–5	$O(kq_i n)$
	Step 6	$O(kQ^3 n^2)$, where $Q = \sum_{i=1}^l q_i$
Step 6	Steps 7–9	$O(l + Q)$
Total complexity		$O(kQ^3 n^2)$

Table 2

Complexity of algorithms A_2 and A_2^B

Step A_2	Step A_2^B	Complexity
Learning: Steps 1–4		$O(q_0 n)$, $O(ln)$, $O(ln)$, $O(1)$ correspondingly
Recognition:		
Step 1	Steps 1–2	$O(kq_{0i} n)$, where $ X_i^0 = q_{0i}$
Step 2	Step 3	$O(q_{0i})$
	Steps 4–5	$O(l)$

End of the table 2

Step A_2	Step A_2^B	Complexity
Steps 1–3	Steps 6–7	Steps 2–5 are executed not more than ln times
Total complexity		$O(kq_0 n^2)$

Complexity of algorithms A_C and A_C^B

Step A_C	Step A_C^B	Complexity
Step 1	Steps 1–3	$O(kq_i^3n)$
Step 2	Steps 4–5	$O(kq_0in)$
Step 3	Step 6–7	Steps 2–5 are executed not more than ln times
Total complexity		$O(k(Q^3 + q_0)n^2)$. If $Q \gg q_0$, then $O(kQ^3n^2)$

It is easy to see that k , n and q_0 are fixed. In the worst case A_1 and A_C can build all set X , therefore it is necessary to reduce quantity of objects processed by these algorithms. Complexity of algorithms is suggested to be reduced by decrease of Q . This is the main objective of this paper and the fundamental idea of described modifications in the class of inductive resolution algorithms.

2. Modification of algorithm A_1 and its properties

The suggested modification of algorithm A_1 is based on the following idea: let us consider features one after another and after consideration of each feature delete objects which can't be used for derivation of object x from X^0 . For this purpose let us introduce function of closeness of two objects by feature s_j and function of closeness of an object to a set of objects by feature s_j :

$$\mu_j^1(x, p) = \begin{cases} 1, & D_j^x \subset D_j^p \\ 0, & D_j^x \not\subset D_j^p \end{cases}; \mu_j^1(x, Y) = \max_{p \in Y} \{\mu_j(x, p)\},$$

where $Y \subset X$. Let $L = \{1, \dots, l\}$ be set of class numbers divided into l subsets containing this numbers, $X_i^0 = X_i \cap X^0$. Function μ_j^1 is used as upper bound of belonging of object x to current class. Let us consider the following algorithm A_1^B :

Step 1. Let $Y_i = X_i^0$, $L = \{1, \dots, l\}$. For all n features execute steps 2-8.

Step 2. Choose feature s_j . For each class X_i , where $i \in L$, execute steps 3-5.

Step 3. Choose from Y_i a pair of objects (p_1, p_2) . If all pairs are considered, return to step 2.

Step 4. Calculate $r = Or_j(p_1, p_2)$. If $\exists j D_j^r = \emptyset$, return to step 3.

Step 5. If $r \notin Y_i$, add r to Y_i : $Y_i := Y_i \cup \{r\}$. Return to step 3.

Step 6. Calculate $\mu_j^1(x, Y_i)$ for all $i \in L$.

Step 7. Delete from L numbers of classes i , for which $\mu_j^1(x, Y_i) = 0$.

Step 8. Delete from Y_i objects p , for which $\mu_j^1(x, p) = 0$.

Step 9. Stop.

Results of A_1^B are interpreted as follows: if $L = \emptyset$ after A_1^B has stopped, then one can't make any conclusions about belonging of x ; if $\exists i \in L$, then $x \in X_i$. Let us show that A_1^B builds those and only those objects which can be logically derived from X_i^0 . For this purpose, let us show first that deletion of classes and objects at steps 7-8 of A_1^B is correct: if x can be derived from X^0 and $\mu_j^1(x, p) = 0$ after consideration of s_j , then afterwards p can't participate in derivation of x , hence deletion of p from Y_i won't lead to not building x by A_1^B .

Theorem 1. Let $\mu_j^1(x, p_1) = \mu_j^1(x, p_2)$, where $p_1, p_2 \in X_i$. If $h \neq j$ or $\mu_j^1(x, p_1) = \mu_j^1(x, p_2) = 1$, then $\mu_j^1(x, Or_h(p_1, p_2)) = \mu_j^1(x, p_1) = \mu_j^1(x, p_2)$.

Let us show that order of consideration of features in algorithm A_1^B doesn't influence the result.

Theorem 2. Let object r be built by algorithm A_1^B from $p_1, p_2, p_3, p_4 \in Y_i$ after consecutive consideration of features s_{j_1} and s_{j_2} . Then r is also built after their inverse consideration.

Finally, let us show that after algorithm A_1^B has stopped, $x \in Y_i$ if and only if x can be logically derived from X_i^0 . It is easy to see that if $Norm(x) \subset Norm(p)$ and $p \in Y_i$, then $x \in Y_i$. Therefore, if p can be derived from X_i^0 and $Norm(x) \subset Norm(p)$, then x can be derived from X_i^0 .

Theorem 3. Object x can be derived from X_i^0 if and only if, after A_1^B has stopped, $i \in L$.

The correctness of algorithm A_1^B follows from theorem 3.

3. Modification of algorithms A_2 and A_C

Let us describe modification of A_2 , which is based on estimation of closeness of objects by partial set of features. Features are also considered one after another. Let ρ_j denote the function:

$$\rho_j(D_h^x, D_h^p) = \begin{cases} (-1)^t, & h \in \{s_1, \dots, s_j\} \\ 1, & h \in \{s_{j+1}, \dots, s_n\} \end{cases}, \quad t = \begin{cases} 1, & D_j^x \neq D_j^p \\ 2, & D_j^x = D_j^p \end{cases}.$$

The value of ρ_j reflect closeness of first j features of objects x and p . Let us also introduce function μ_j^2 of upper bound of partial closeness of objects by j features:

$$\mu_j^2(x, p) = \max\{0, (\sum_{h \in S} a_{ij})^{-1} \times \sum_{h \in S} a_{ih} \rho_j(D_h^x, D_h^p)\},$$

where $a_{ij} \in R$, $\forall i, j$ $a_{ij} \geq 0$, $\forall i$ $\sum_j a_{ij} > 0$, $\|a_{ij}\|$ is a matrix corresponding to set $\{1, \dots, l\} \times |S|$. If $h=0$,

then $\mu_j^2(x, p) = 1$. Let us consider the modification of A_2 , denoted by A_2^B : L is divided into l subsets containing class numbers; μ_j^2 is used as upper bound of belonging of x to a class.

Step 1. Let $\forall i = \overline{1, l}$ $j_i = 0$. Choose arbitrary class X_i .

Step 2. For all $x^0 \in X_i^0$ calculate $\mu_{j_i}^2(x, x^0)$.

Step 3. For all $i = \overline{1, l}$ calculate $P_i(x) = \max_{x^0 \in X_i^0} \{\mu_{j_i}^2(x, x^0)\}$.

Step 4. Calculate $P(x) = \max_i \{P_i(x)\}$.

Step 5. Let $P_{i_0}(x) = \max_i \{P_i(x)\}$. If $j_{i_0} = n$, then go to step 7.

Step 6. Go to next feature in X_{i_0} : $j_{i_0} := j_{i_0} + 1$. Return to step 2.

Step 7. Include x in class X_{i_0} . Stop.

Finally, let us consider the modification of A_C , which combines A_1^B and A_2^B , denoted by A_C^B :

Step 1. Let $L = \{1, \dots, l\}$, $\forall i = \overline{1, l}$ $j_i = 0$, $Y_i = X_i^0$. Choose arbitrary class X_i .

Step 2. If $i \in L$, then execute steps 3-8 of A_1^B for X_i and feature j_i , otherwise go to step 4.

Step 3. If $\exists p \in Y_i$ $Norm(x) \subset Norm(p)$, then denote $i_0 = i$ and go to step 7.

Step 4. Execute steps 2-4 of algorithm A_2^B .

Steps 5-7 of algorithm A_C^B are equal to steps 5-7 of algorithm A_2^B .

As opposed to A_1^B , A_2^B calculates belonging of any $x \in X$. Thus, by using of algorithm A_C^B the belonging of any $x \in X$ can also be determined.

4. Complexity of modified algorithms

In the worst case complexity of A_1^B , A_2^B and A_C^B is equal to the complexity of A_1 , A_2 and A_C correspondingly. Calculation of complexity of algorithms is described in details in tables 1-3.

Let us now estimate the change of Q in A_1^B and A_C^B in comparison with A_1 and A_C correspondingly. Let Y_i be a set of objects before consideration of feature s_j ; Y_i^1 and Y_i^B are sets built as a result of consideration of s_j by A_1 and A_1^B correspondingly; $|Y_i^1| = q_i^1$, $|Y_i^B| = q_i^B$. It is necessary to compare q_i^1 and q_i^B . Let us assume that in Y_i all feature values are equifrequent. For simplification reasons, we won't delete equal and non-existent objects from Y_i^1 and Y_i^B . In A_1 , set Y_i^1 consists of all resolvents built from Y_i using s_j . In A_1^B , set Y_i^B consists of only those objects p , for which $\mu_j(x, p) = 1$: $Y_i^B = \{p \mid d_{jt}^x \leq d_{jt}^p, t = \overline{1, k}\}$. Therefore, if $d_{jt}^x = 1$, then $d_{jt}^p = 1$, hence if $c(D_j^x)$ contains m_j ones, then in average $q_i^B = 2^{-m_j} q_i^1$. Thus, if $c(x)$ contains $m = \sum_{j=1}^n m_j$ ones, then $Q^B = 2^{-m} Q^1$, where Q^1 and Q^B are total quantities of objects built by A_1 and A_1^B . Hence in average A_1^B works 2^{3m} times quicker as A_1 , and therefore A_C^B works 2^{3m} times quicker as A_C .

Conclusion

Three algorithms for solution of pattern recognition problem are considered. They are modifications of algorithms described in [1] and use branch and bound method. Computational complexities of both unmodified and proposed algorithms are calculated. The running time of modified algorithms is shown to depend exponentially on the structure of given information.

References

1. Shut, O.V. Synthesis of Algorithms for the Solution of Pattern Recognition Problems in Finite-Dimensional Discrete Spaces / O.V. Shut // Vestnik BSU. – 2014. – Series 1; № 1 (in press).
2. Krasnoproshin, V.V. Recognition with Learning as a Problem of Choice / V.V. Krasnoproshin, V.A. Obratsov // Digital Images Processing. – Minsk, 1998. – P. 80–94.
3. Ablameyko, S.V. Pattern Recognition and Image Analysis: Theory and Experience of Solution of Practical Problems / S.V. Ablameyko, V.V. Krasnoproshin, V.A. Obratsov // International Congress of Computer Science: Information systems and Technologies. – Minsk : BSU, 2013. – P. 434–444.
4. Gafurov, S.V. The Problem of Incompleteness of Information in Pattern Recognition / S.V. Gafurov, V.V. Krasnoproshin, V.A. Obratsov // Vestnik BSU. – 2007. – Series 1; № 2. – P. 113–116.

THE MODIFIED PRINCIPAL COMPONENT ANALYSIS OF INFORMATION ENCRYPTED USING DETERMINISTIC CHAOS

A.V. Sidorenko, I.V. Shakinko
Belarusian State University, Minsk
e-mail: SidorenkoA@yandex.ru

Modification of the principal component analysis is proposed. The results obtained by the singular spectrum analysis and by the proposed method for information encrypted using deterministic chaos are compared.

Introduction

As modern information technologies are more and more widely used in telecommunication systems, information security plays an increasingly important part. The use of deterministic chaos in encryption imposes new requirements to ensure the cipher strength. In the encryption algorithm based on deterministic chaos it is necessary to control the chaotic state of output sequences. Sometimes the available methods to cope with this task are inadequate. Given that the output sequences of the encryption algorithms based on deterministic chaos are not stationary, singular spectral analysis (SSA) can be used for their processing [1]. Since periodic processes in these realizations are nonstationary too, the application of SSA is insufficient to detect the deterministic components [2]. This paper suggests modification of the principal component analysis; the results obtained by the proposed method are analyzed.

1. Algorithm of the modified principal component analysis

An algorithm of the modified principal component analysis comprises the following steps:

- clumpiness test for a system of iterated functions;
- formation of the point allocation using one coordinate;
- conversion of the point allocation into the matrix;
- application of the principal component analysis.

Let us consider these steps in detail.

1. Clumpiness test for a system of iterated functions [3].

For definiteness, we consider the time series A containing N samples

$$A = \{a_1, a_2, \dots, a_{N-1}, a_N\} \quad (1)$$

The range d of the time series values is calculated as follows:

$$d = a_{\max} - a_{\min}, \quad (2)$$

where a_{\max} and a_{\min} – maximal and minimal values taken by samples of the time series.

The resulting range is partitioned into four equal intervals. A square field is drawn on the plane. The corners of the square field are numbered from the lower left one, moving counterclockwise. The first point is marked at the center of the square. Next, a new point is marked in the middle of the segment connecting the center point and the corner with a number identical to that of the interval, to which the value of the first sample belongs. Then the procedure is repeated for the following samples with the only difference that the interval is measured from the last point marked. Thus, the point allocation is formed in the square.

2. Formation of the point allocation using one coordinate.

If the coordinate axes OX and OY are along the two perpendicular sides of the square, each point belonging to the square can be associated with two numbers: coordinate x and coordinate y . Considering that each sample of the initial time series corresponds to a certain point, it is possible to form a set of the pairs “sample number – point coordinate”. Coordinates x or y for all pairs are selected as a “point coordinate”. Each pair of values is regarded as the coordinates of some point belonging to a new point allocation in the rectangular area B .

3. Conversion of the point allocation into the matrix.

The rectangular area B is partitioned into R equal, non-overlapping parts as

$$m \cdot n = R, \quad (3)$$

where m – number of horizontal divisions , n – number of vertical divisions. Let us define the matrix C containing n rows and m columns, the elements of which are found in the following way:

$$c_{ij} = N_{ij}, \quad (4)$$

where N_{ij} – number of the points belonging to ij – section of the region B , $i = 1..n$, $j = 1..m$.

4. The matrix C is converted by the principal component analysis that includes the following procedures:

- centering and normalizing

$$c_{ij}^* = \frac{(c_{ij} - \bar{c}_j)}{s_j}, \quad (5)$$

where

$$\bar{c}_j = \frac{\left(\sum_{i=1}^n c_{ij} \right)}{n}, \quad (6)$$

$$s_j = \frac{\sqrt{\left(\sum_{i=1}^n (c_{ij} - \bar{c}_j)^2 \right)}}{n}, \quad (7)$$

$i = 1..n$, $j = 1..m$;
– singular value decomposition (SVD)

$$C^* = USV^t, \quad (8)$$

where t – symbol of the matrix transposition;

- calculation of the score matrix

$$T = US; \quad (9)$$

- calculation of eigenvalues

$$L = T^t T; \quad (10)$$

– normalization of each eigenvalue to the sum of all eigenvalues (determination of a level for the contribution of each component):

$$\lambda_i = \frac{l_i}{\sum_{j=1}^k l_j}, \quad (11)$$

where

$$k = \min(n, m), \quad (12)$$

$i = 1..k$.

2. Modified principal component analysis of information encrypted using dynamic chaos

The messages processed were 22000 samples in length. The scheme of the encryption algorithm based on the Feistel cipher and deterministic chaos was used. Four encryption modes were analyzed: cipher-block chaining (CBC), cipher feed-back (CFB), electronic codebook (ECB), and output feedback (OFB). The iteration number z of the chaotic map varied from 1 to 1024. Sizes of the formed matrices were chosen equal: $m = 100$, $n = 100$, and $m = 1000$, $n = 1000$. The following methods were used for graphical representation of the formed matrices. Each element of the matrix, depending on its value, was assigned to the specific color: minimal value of the matrix – white, maximal value – black. A larger value of the element corresponds to a greater shift to black color. The resultant graphical matrix representations are given in fig. 1.

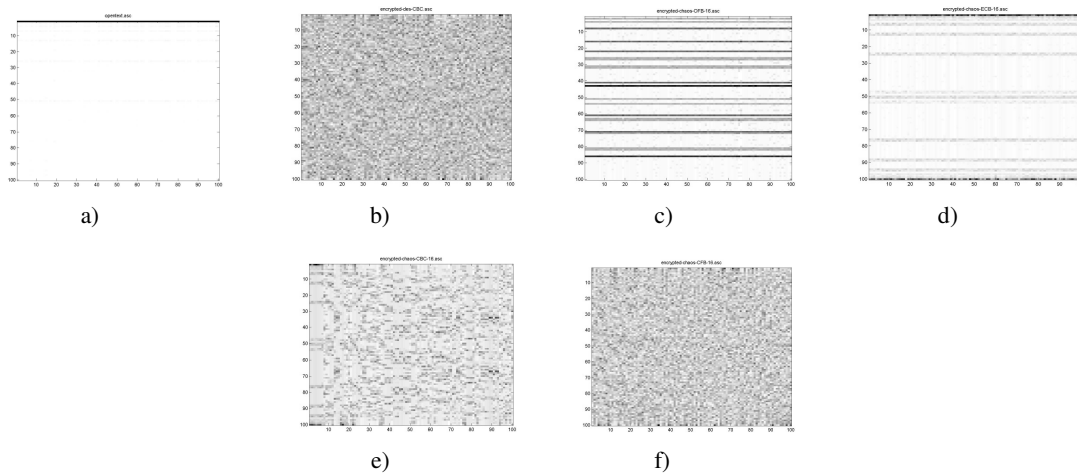


Fig. 1. The graphical matrix representations ($m = 100, n = 100$) obtained for the plaintext (a); algorithm *des* (b); algorithm based on deterministic chaos using OFB (c), ECB (d), CBC (e), and CFB (f) operation modes

Graphical representation for the case of the plaintext of the resultant matrix is a single horizontal black line at the top of the image. When using the encryption algorithm based on deterministic chaos in OFB mode, image is a set of dark, well-defined lines. In the graphical representation corresponding to ECB mode the lines can also be traced but they are rather pale, approaching the background. The exception is the two horizontal lines located at the bottom and the top of the image. For the standard encryption algorithm *des* and the encryption algorithm based on deterministic chaos in CFB mode, the formed images represent the dots uniformly distributed over the entire area. Unlike the encryption algorithm based on the deterministic chaos mode CBC, where unevenness of the point locations is observed, a certain structure is revealed. The graph for a relative-contribution level of the first principal component (r) was plotted (fig. 2) as a function of the iteration number (z) for the chaotic map. A relative-contribution level of the first principal component is a level of the contribution made by the first principal component, normalized to the similar parameter obtained for the plaintext.

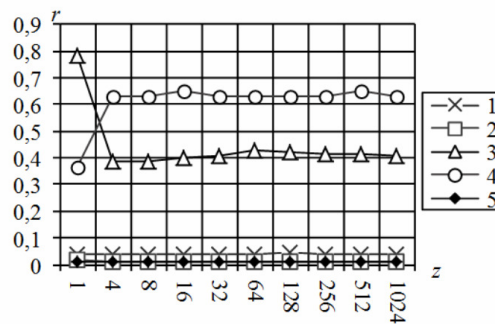


Fig. 2. The graph for a relative contribution level of the first principal component (r) (modified principal component analysis) versus the iteration number (z) of the chaotic map: encryption algorithm based on deterministic chaos using CBC (1), CFB (2), ECB (3), and OFB (4) operation modes; algorithm *des* (5)

For the encryption algorithm based on the deterministic chaos using the modes OFB and ECB, a relative contribution level of the first principal component is about 0.6 and 0.4, respectively, whereas with the use of the encryption modes CBC and CFB this parameter is below 0.1. However, when using the mode CBC, values of a relative contribution level for the first principal component are higher than those obtained for CFB mode and for the standard encryption algorithm by a factor of three and more (fig. 3). A high relative contribution level of the first principal component indicates the presence of determinism in the encrypted messages. Thus, based on the graphical representation of the matrix and on the obtained principal component levels, we can infer that CFB mode of operation is suitable for the encryption algorithm based on the Feistel cipher and deterministic chaos as distinct from CBC, ECB, and OFB modes, the use of which leads to the encrypted messages revealing the deterministic components.

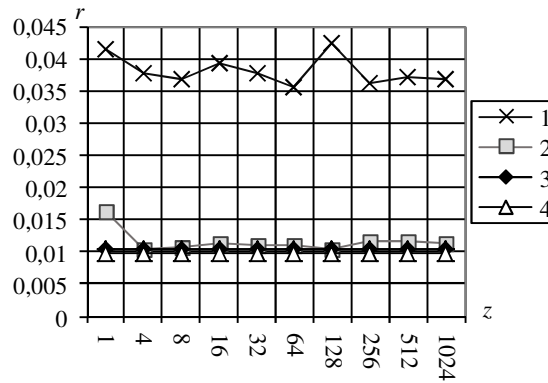


Fig. 3. The graph for a relative contribution level of the first principal component (r) (modified principal component analysis) versus the iteration number (z) of the chaotic map: encryption algorithm based on deterministic chaos using CBC (1) and CFB (2) modes of operation; algorithm *des* (3); algorithm *aes* (4)

3. Comparison of the results obtained by the singular spectrum analysis and proposed method

To process realizations of the encrypted messages by a singular spectrum analysis, the decomposition procedure was conducted for 1000 components; "centering" was carried out. For a visual representation of the data obtained by the singular spectrum analysis, the graph for a relative contribution level of the first principal component (r) versus the iteration number (z) of the chaotic map was constructed (fig. 4).

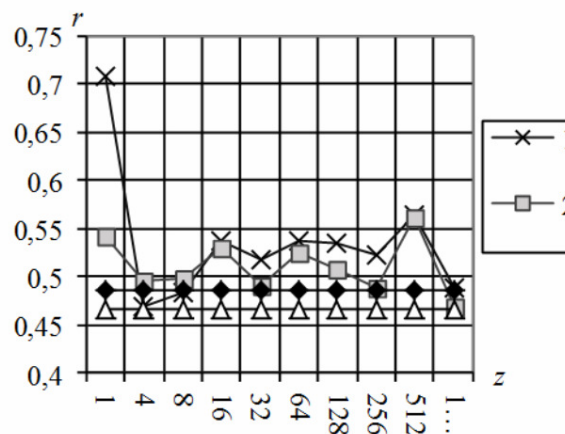


Fig. 4. The graph for a relative contribution level of the first principal component (r) (singular spectrum analysis) versus the iteration number (z) of the chaotic map: encryption algorithm based on deterministic chaos using CBC (1) and CFB (2) modes; algorithm *des* (3); algorithm *aes* (4)

As shown, values of a relative contribution level of the first principal component obtained for the algorithm based on the Feistel cipher and deterministic chaos using CBC and CFB operation modes are approximately coincident, falling within the range from 0.45 to 0.6 (excepting the value obtained for the iteration number that is equal to 1). In this way the results obtained using the singular spectrum analysis, unlike the proposed method, reveal no differences between realizations of the encrypted messages.

Conclusion

Modification of the principal component analysis is proposed that consists in forming the matrix from the initial data with the use of the clumpiness test for systems of iterated functions.

It has been found that the proposed modified analysis of principal components allows one to analyze quantitatively (by a level of the contribution made from the principal components) and qualitatively (in the form of the graphically represented matrices) the output sequence of the encryption algorithm based on the Feistel cipher and deterministic chaos.

The results obtained by the singular spectrum analysis and by the proposed method have been compared.

It has been established that the modified principal component analysis, in contrast to the singular spectrum analysis, offers detection of determinism in the messages encrypted by the encryption algorithm based on the Feistel cipher and deterministic chaos using CBC operation mode.

On the basis of the graphical matrix representations and proceeding from the obtained principal-component levels, it has been found that the mode CFB is suitable for the encryption algorithm based on the Feistel cipher and deterministic chaos as distinct from CBC, ECB and OFB modes, the use of which leads to encrypted messages with the components revealing a high degree of determinism.

References

1. Sidorenko, A.V. Encryption is based on deterministic chaos using the singular spectrum analysis / A.V. Sidorenko, I.V. Shakinko // Proc. of the 69th science conf. of students and postgraduates of BSU / Belarusian State University – Minsk, 2012. – Vol. 1. – P. 276–279.

2. Borog, V.V. Combined method of identifying hidden anomalies in chaotic time processes / V.V. Borog, A.V. Kryanev, D.C. Udumyan // Fundamental physical and mathematical problems and modeling of technical and technological systems: coll. of science proc. – Moscow : Janus-K, 2009. – Vol. 12. – P. 536–546.

3. Meckler, A.A. Application of the apparatus nonlinear dynamical systems analysis for EEG signal processing / A.A. Meckler; ed. by E.V. Kalashnikov // Actual problems of modern mathematics: science notes. – St. Petersburg : Pushkin Leningrad State University, 2004. – Vol. 2. – P. 112–140.

A SCHEME FOR NODULES DETECTION ON CT IMAGES OF LUNGS

E. Snezhko¹, S. Kharuzhyk², A. Tuzikov¹, V. Kovalev¹, I. Safonau¹

¹United Institute of Informatics Problems of the NAS of Belarus, Minsk;

²N.N. Alexandrov National Cancer Centre of Belarus, Lesnoy

e-mail: eduard.snezhko@gmail.com

The results of development and application of the software for computer-assisted lung cancer diagnosis are presented. The software is intended to effectively localize pulmonary nodules on computed tomography images of lungs and consists of several stages: lungs segmentation, nodule candidates selection and nodule candidates filtering.

Introduction

According to the World Health Organization (WHO) lung cancer remains the leading cause of death of men among all malignant tumors. The expected 5-years survival rate is about 15%, which is not much greater than 6% in 1960th. One of the reason of such a statistics is the fact that lung cancer is hardly diagnosed on the yearly stages when it is almost asymptomatic. Computed tomography is the primary modality for imaging lung cancer and Computer-Aided Diagnosis (CADx) becomes more valuable tool in medical practice [1].

The purpose of this particular paper is to present results of an experimental study of the CADx software developed in framework of Belarussian national programme. One of the main purpose of this software was the localization of pulmonary nodules in lungs on computed tomography images (CT scans): solitary pulmonary nodules (SPN) and multiple nodules as well. The efficiency of nodules localization was intended to be as high as the level of the best practice.

1. Materials

In this study we used CT images of lungs of 87 patients with lung nodules collected at the N.N. Alexandrov National Cancer Centre of Belarus. CT scanning was performed on a multi-slice scanners with the standard kV and mA settings during the one-breath hold. The voxel size of tomograms was in the range of 0.65-0.74 mm in the axial image plane with the slice thickness equal to the inter-slice distance varied from 1.25 mm to 3 mm. CT images contained from 104 to 407 slices. Target size of nodules varied from 4 mm to 56 mm. The expert radiologist found 173 nodules on these 87 CT scans. Solitary pulmonary nodules were present on 82 CT scans. Multiple nodules (from 7 up to 21) were present on 5 CT scans. No intravenous contrast agent was administered before the collection of scan data.

2. Methods

The process of nodules detection utilized during development of the CADx software in general conforms to the “classical” scheme and contains the following stages: 1) lungs segmentation on CT images; 2) nodule candidates selection; 3) nodule candidates filtering based on a set of decision rules.

2.1. Lungs segmentation

A number of modern lung image processing techniques of in practice utilize either manual segmentation or process the whole image extent without separation lungs from the outer tissues and other objects. Since CADx software for the particular task of nodules detection in lungs implies the processing and analysis of considerably large volume of radiological images, fully automatic lungs segmentation is the prerequisite of all the process.

Stages of lungs segmentation was were described in paper [2]. The first step is to apply the adaptive thresholding to the whole 3D CT image. Adaptive thresholding allow computing the level of thresholding by means of iterative process. Let T^i to be the thresholding level at the iteration i , and let μ_b and μ_n are the mean values in Hounsfield units (HU) of the “inner” part (voxels with higher or equal values than T^i) and “outer” part (voxel with lower values than T^i) of the image. Thus at the iteration $i + 1$ threshold is calculated as $T^{i+1} = \frac{\mu_b + \mu_n}{2}$. This procedure stops at the iteration c when $T^c = T^{c-1}$. The

initial value of T^0 was selected to be equal 10 HU. After termination of the iterative process “outer” we acquired a 3D image mask with values 0 in the voxels which correspond to “outer” voxels of a CT image and 1 – to “inner” voxels the image. The example of an initial CT image and the resulting image mask calculated as a result of adaptive thresholding is depicted on fig. 1.

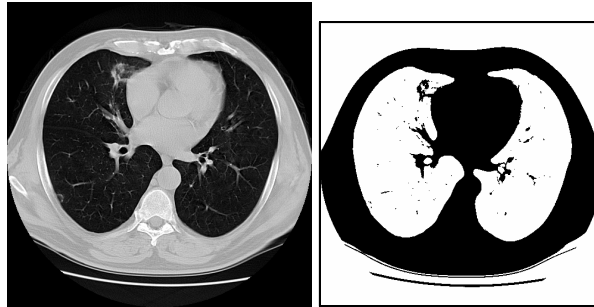


Fig. 1. Initial image (left) and the resulting image mask (right) after adaptive thresholding of the initial image

On the next step we invert values of the mask (*ones to zeros and vica versa*, fig. 2, left) obtained after thresholding and fill holes of the mask. As a result of this operation the mask will contain several connected components, including the largest one corresponding to body of a patient (fig. 2, center). The other connected components which correspond to the table or (perhaps) other objects are removed on the next step of the segmentation process, when we extract only the largest component (fig. 2, right).

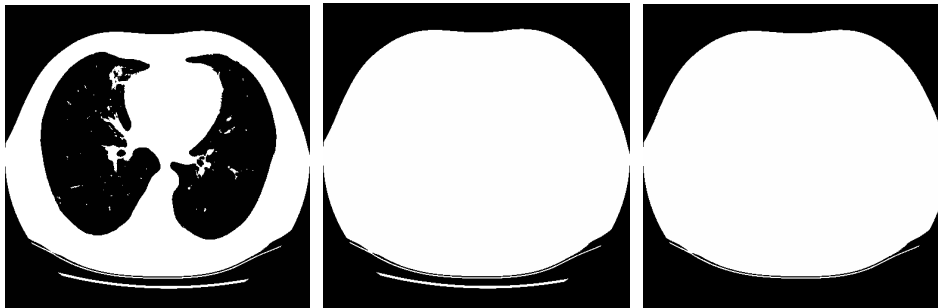


Fig. 2. Extraction of body mask which on a CT image

Applying AND logical operation to the inverted mask from fig. 2, right and the mask from fig. 1, right, one obtain the mask which corresponds only to the lungs on the initial CT image (fig. 3, left). However, this mask contains a lot of small holes, scratches and other defects, so a number of nodule candidates may fall out of consideration if we apply this mask to the initial CT image and consider only the voxels corresponding to *ones* of the mask obtained till this moment. To solve the issue mathematical morphology operations “opening” and “closing” should be applied to the mask. The resulting mask is depicted on fig. 3, center, and the lungs segmented on the initial CT image is shown on fig. 3, right.

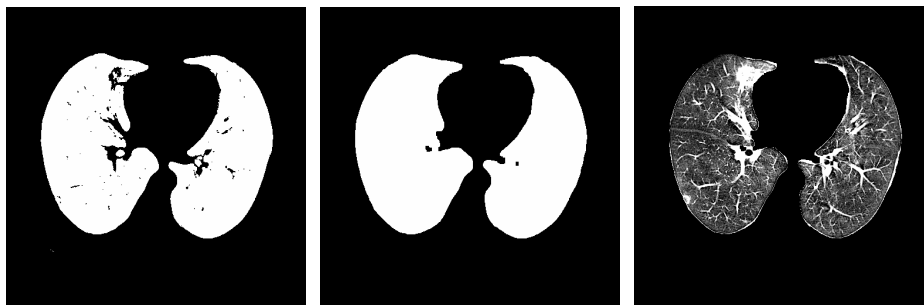


Fig. 3. Draft lungs mask (left); smoothed lungs mask (center); segmented lungs on a CT image (right)

2.2. Nodule candidates selection

Since “large” and “small” nodules manifest themselves on CT images a bit different way, all the following stages of nodule candidates selection and further filtering is performed in two passes with different presets.

The initial nodule candidates selection is performed by means of thresholding of the segmented lungs followed by morphological “opening” and “closing” operations. For “small” nodules thresholding level was selected equal to -400 HU, and radius of the structuring element for morphological operations equal to 2 voxels. The corresponding presets for “large” nodules were equal to -150 HU and 3 voxels. Then connected components are extracted and labeled by their index numbers. All the connected components are assumed as nodule candidates at this stage, so called “blobs”. The following geometry and intensity parameters are calculated for each candidate: volume (number of voxels related to blob); minimum and maximum coordinates of voxels along all three axis X, Y and Z; coordinates of blob center; lengths of principal axis, e.g. radii of inertia ellipsoid; ratio of minimum and maximum radii of inertia ellipsoid; ratio of blob volume to a bounding volume; number of voxels on the surface of a blob (surface area); ratio of surface area to the blob volume; average voxel intensity inside every blob; variance of voxel intensity inside every blob. All these parameters are necessary for further nodule candidates filtering.

2.3. Nodule candidates filtering

The procedure of nodule candidates filtering is applied to the list of blobs, each having a number of precalculated geometry and intensity parameters. Filtering is based on a set of rules, which define thresholds for blob parameters. Thresholds were chosen differently for “small” and “large” nodule candidates. Thresholds relate to the following: lower blob volume limit; lower blob lengths in axial plane; lower blob depth limit; lower limit for ratio of blob volume to the volume of inertia ellipsoid estimated; lower limit for ratio of blob surface area to blob volume; lower limit for ratio of minimum to maximum radii of inertia ellipsoid; lower limit for average intensities of voxels related to a blob.

It is sufficient for any parameter of any given nodule candidate not satisfy any threshold mentioned above for removing this blob from consideration. Usually a couple dozen nodule candidates of initially a couple of hundred remain after this step.

Then region growing algorithm is applied for each blob left till the moment. This is done to ensure that a part of vessels or other large object with irregular structure will not pass the filtering stage. Additional verification of grown blobs is performed. If volume of some blob was increased more than 2.5 times, this blob is removed from consideration. Finally, we remove duplicates from the list of blobs, i.e. if any two blobs intersect we remove the smallest one.

After blobs filtering we receive a list of blobs with detailed information about them (voxel coordinates, geometry and intensity parameters) which could be used further at visualization and analysis stages.

3. Results

The algorithm on pulmonary nodules detection was tested on the set of 87 CT scans. The expert radiologist reviewed all the images carefully in advance and detected 173 nodules. The proposed algorithm detected 123 true positive (TP) and 179 false positives (FP) findings. Thus, sensitivity (recall) of the nodule detection algorithm was $TPR = TP / (TP + FN) = 123 / 173 = 71.1\%$. Ratio of FP and TP findings was equal to $179 / 123 = 1.46$. The most often reason for false positive findings were joins between nodules and pleura, mediastinal and peripheral vessels, mediastinal contour, and septum between lungs. At the same time, the algorithm found additional nodules on 10 CT scans which were not detected by the expert radiologist. Processing of one CT scan took from 110 to 350 seconds depending on number of axial slices; average processing time was equal to 3.25 minutes.

Due to significant difference in nodule number from one CT scan to another (it could be from 1 up to 21), two groups of CT scans were considered: scans with small amount of nodules (from 1 to 4) and multiple nodules (from 5 and more).

82 CT scans were related to the first group. The expert radiologist detected 115 nodules on these scans, and the proposed algorithm detected 89 TP findings and 177 FP findings. Sensitivity of the algorithm was $TPR = 77.4\%$. Ratio of FP to TP findings was equal to 1.98. Of these examinations there was a group of 60 scans with solitary (single) pulmonary nodules. The statistics for these scans is as follows: $TPR = 90.0\%$, $FP/TP = 2.31$.

Multiple nodules were present on 5 CT scans: 5, 6, 7, 19 and 21 nodules respectively (see Figure 4 for example), 58 in total. The proposed algorithm detected 34 TP and 2 FP findings, i.e. sensitivity on this group of scans was equal to 58.6% and ratio of FP to TP findings was equal to 0.06.

Conclusion

Results obtained with this experimental study on 87 CT scans of lungs allow to draw the conclusions that the proposed algorithm of nodules detection, which consists of lungs segmentation, nodule candidates selection and further filtering steps may be used as a effective tool for early lung cancer diagnosis as well as for training. The algorithm works better for solitary pulmonary nodules detection (TPR = 90.0%). If we consider CT scans with more than one nodule (up to 4) TPR decreased to 77.4%, and for CT scans with multiple nodules (more than 5) TPR was equal to 58.6%. However, it should be noted that for multiple nodules the algorithm may significantly facilitate reading of CT scans in order not to miss true positive findings with very small percentage of false positive findings. Also, it was shown that the algorithm found additional nodules on 10 CT scans which were not detected by the expert radiologist.

This work was done in framework on the National project IT11/4-04.

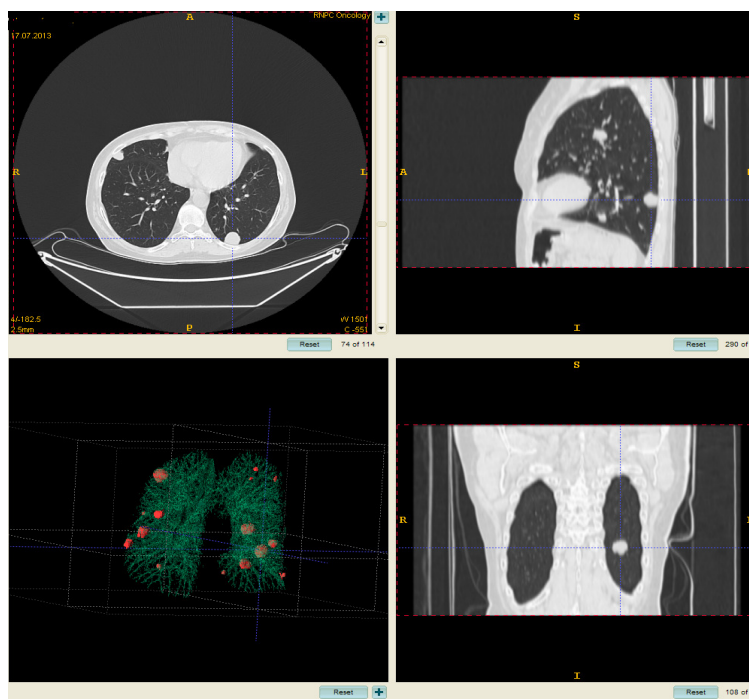


Fig. 4. Example of multiple nodule detection on a CT scan using the algorithm proposed

References

1. Computer-assisted diagnosis of nodular lesions in the lung // S.A. Kharuzhyk [et al.] // *Voprosi Onkologii*. – 2011. – Vol. 57 (1). – P. 25–35.
2. Automatic Lung Segmentation for Accurate Quantitation of Volumetric X-ray CT Images / S.Hu [et al.] // *IEEE Transactions on Medical Imaging*. – 2001. – Vol. 20, no. 6. – P. 490–498.

IMAGE PROCESSING FOR FUNDUS IMAGE ANALYSIS

V. Starovoitov

United Institute of Informatics Problems of the NAS of Belarus, Minsk

e-mail: valerys @newman.bas-net.by

Problems of automated fundus image processing and analysis are discussed in the paper. Procedures for retinal image enhancement are selected and tested. The problem of automatic quality assessment of retinal images is discussed also.

Introduction

Diabetic retinopathy is the leading cause of blindness in the world. The World Health Organization estimates that 135 million people have diabetes mellitus worldwide and that the number of people with diabetes will increase to 300 million by the year 2025 [1]. Timely detection and treatment for diabetic retinopathy prevents severe visual loss in up to 50% of the patients. Therefore, an automatic or semi-automatic system able to detect various type of retinopathy is a vital necessity to save many sight-years in the population.

1. Retina and its features in a digital image

The fundus of the eye is the interior surface of the eye, opposite the lens. Retina is a part of fundus. It is light-sensitive layer of tissue, lining the inner surface of the eye. In the center of the retina is the optic nerve, a circular to oval white area measuring about 2x1.5 mm across. From the center of the optic nerve radiates the major blood vessels of the retina. It is approximately 17° (4.5-5 mm), or two and half disc diameters from the disc. It can be seen the slightly oval-shaped, blood vessel-free reddish spot. The fovea is at the center of the area known as the macula by ophthalmologists. The fovea is less than 1 mm in diameter (fig. 1). A circular field of approximately 6 mm around the fovea is considered the central retina while beyond this is peripheral retina stretching to the ora serrata, 21 mm from the center of the retina (fovea). The total retina is a circular disc of between 30 and 40 mm in diameter [2].

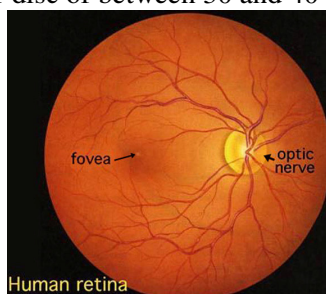


Fig. 1. A typical image of a human retina with indication of two basic objects: optic disk (OD) and fovea

In a digital image retina is presented in the RGB color model and all the metric relations mentioned above are measured in pixels and depend on the camera field of view (FOV). Usually images produced by a fundus camera has a circular shape, but retina may be differently presented in the picture and the image size may vary (fig. 2).

Fundus usually has a reddish-orange color. If it reflects a disease, it changes color and pigmentation (fig. 3). Color may vary also due to optical distortion in image registration (fig. 4).

Images may be investigated immediately after registration (on-line analysis) or stored for the later analysis (off-line analysis). For both variants image quality is important. In the first case, image quality may be evaluated interactively, but in the second one it is desirable to perform this automatically.

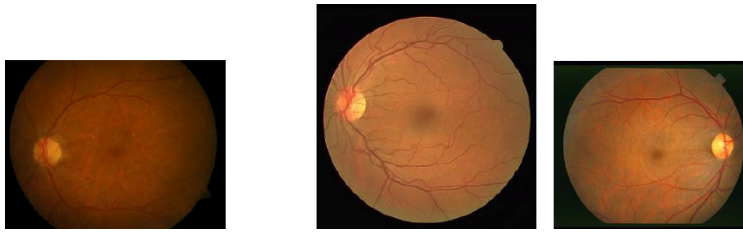


Fig. 2. Sample images from databases: left, DIARETDB1 (FOV=50°, image size 1500x1152 pixels); center, DRIVE (FOV=45°, size 565x584); right - STARE (FOV=35°, size 700x605)

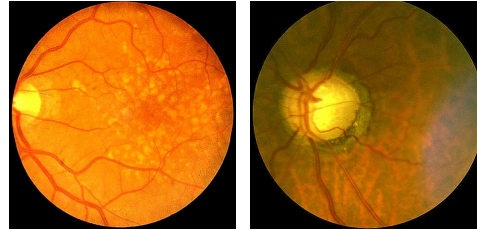


Fig. 3. Images of unhealthy retina: age-related macular degeneration (left, macula is not uniform) and advanced glaucoma (right, optic disk border is not uniform)

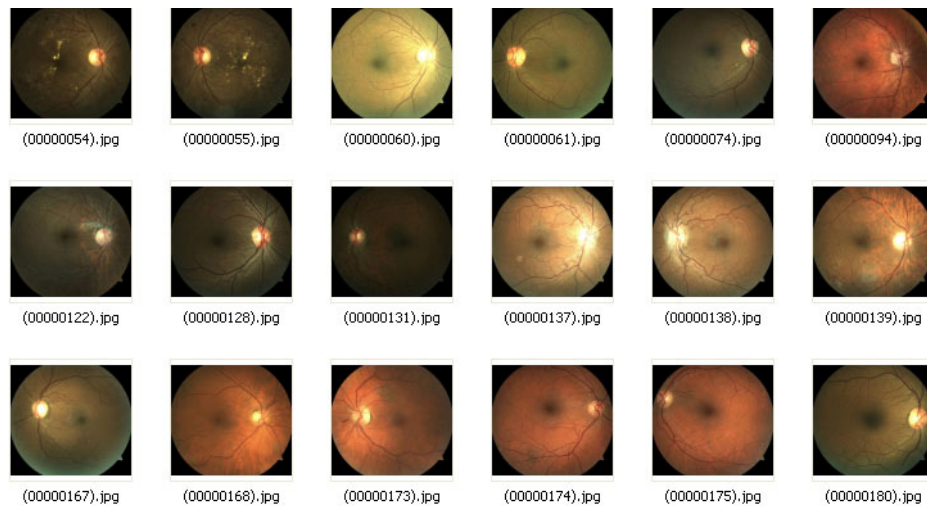


Fig. 4. A publicly available image database DMED is a collection of 169 images of different color and contrast

2. Diabetic retinopathy

Various systems for automatic or semi-automatic detection of retinopathy with digital fundus images have been developed [3], but the problem is not solved. The results obtained are promising but the initial image quality is a limiting factor; this is especially true if the machine operator is not a trained photographer. An accurate quality assessment algorithm can allow operators to avoid poor images by re-taking the fundus image, eliminating the need for correction algorithms. In addition, a quality metric would permit the automatic submission of only the best images if many are available.

Many researchers study image quality methods, and in particular retinal image quality. Investigations in this area started just 10-15 years ago. By now there is no a unique approach to the actual practical problem how to evaluate in real time quality of just registered retinal image until a patient is sitting in the clear in the medical cabinet. Images with adequate quality may be stored, sent for consultation to a medical expert to other clinic or even country by telemedicine.

Several databases containing retinal images and used in the scientific literature for result comparison were collected from Internet. Literature review of state-of-art papers connected to the studied problem was done. Some experiments described in the papers; tested some own ideas.

3. Retinal image processing

Analysis of the literature and publicly available image database shows that the most important procedures for retinal image processing are the following: contrast enhancement, color enhancement and segmentation. While segmentation is object oriented for: 1) background separation, 2) optic disk segmentation, 3) macula segmentation, and 4) vessel segmentation. The very first step of retinal image processing is fast evaluation does the image contain any retina or clutter (fig. 5).



Fig. 5. Nonretinal images registered by fundus cameras

Background in fundus digital images is not black and uniform often. If it is dark enough, it may be segmented by a simple global thresholding algorithm. In opposite case (fig. 6) other method must be used. In fig. 5 one can see that there are two horizontal black stripes in the image: in the top and in the bottom, but the rest of background is dark-green.



Fig. 6. im0055 image from STARE database with nonuniform illumination and low color difference between retina and background

Segmentation of background is important for saving it black after the following retina color transformations and no using this area during analysis of the real retinal area.

A simple way for retina boundary detection and background segmentation is based on the observation that four image corners and the adjacent regions always belong to background. Using the fact we segment the original image into object and background, for example by the Otsy method (fig. 7).

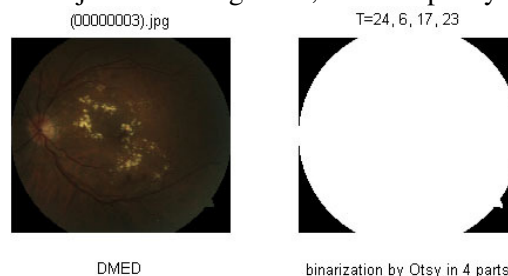


Fig. 7. Retina mask construction by image separation into 4 parts and global thresholding. Threshold values are shown in the picture

Several variants of color enhancement were tested, three of them are presented in fig. 8 and 9.

4. Quality assessment in fundus images

The main problem in retinal image quality evaluation is lack of formal medical protocols (national or international) describing which images are of good quality, which are bad. In this situation IT specialists develop different quality measures of digital images itselfs. The quality assessment measure for this type of images must be *non-reference*. Despite of the importance of this problem it is still a widely neglected field of research especially with regard to fundus imaging [4].

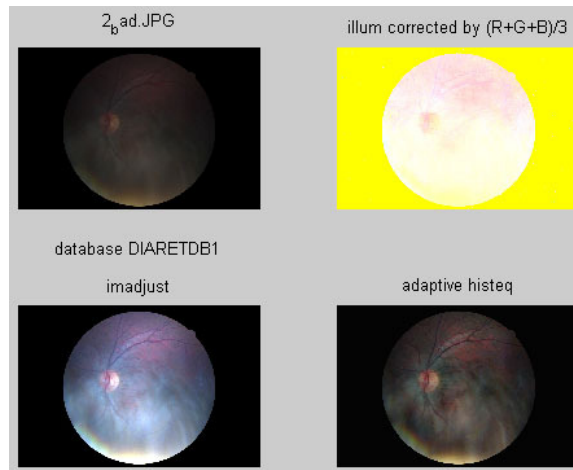


Fig. 8. Illumination correction experiments with images from DIARETDB1 retina database

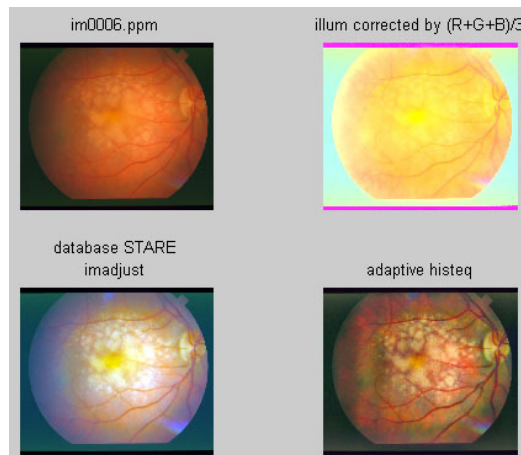


Fig. 9. Illumination correction experiments with images from STARE retina database

There are only several publications dealing with retinal image quality assessment. They may be divided into two classes:

(i) Segmentation based approaches detect anatomical structures, while it is assumed that the segmentation will fail on low quality images due to the bad recognizability. Fleming et al. [5] measure the quality by evaluating the vessel tree in the regions around the macula and the optic nerve head. Giancardo et al. [6] measure the densities of vessels for different regions in the image. The vessel densities and a 5-bin-histogram of each color channel are used as features for classification. In [10] three properties of the optic nerve in order to locate it are used: 1) the optic nerve normally appears as a bright disk approximately $1500\mu\text{m}$ in diameter, 2) large vertical blood vessels enter the nerve from above and below, and 3) blood vessels in the retina tend to converge at the nerve. The nerve location algorithm forms three images, each indicating the strength of one of the three properties, and computes a weighted average of the three images. The position of the pixel of maximum intensity in the average image indicates the position of the optic nerve.

The fovea can be identified in the blue plane image. The fovea is located 4.5mm temporal to the optic nerve and is marked by yellow pigment in the retina, which shows as a dark spot in the blue plane image.

(ii) Histogram based approaches use information gained by image statistics to identify low quality photos. Lalonde et al. [7] evaluate the histogram of an input image's gradient magnitude and local histogram information of its gray values. Reference histograms are calculated out of images showing good quality and compared with the input image's histograms for classification. Lee et al. [8] compute a quality index by convolving the intensity histogram of the input image with the template intensity histogram from good retinal images. Image Structure Clustering from [9] characterizes the image quality by distribution of image intensities itself and the ability to cluster the image into contained anatomical structures.

Some semiformal recommendation of retinal quality grades were done by The National Screening Committee UK in 2012. Images must be utilized only if the grader is confident the quality is sufficient. 3 categories of retinal image quality were defined: good, adequate, inadequate.

Good macula and disc images (the mentioned objects are in the image center) when: 1) center of fovea (or disc) < 1DD from center of image (where DD – disc diameter); 2) vessels clearly visible within 1DD of center of fovea (or disk); 3) vessels visible across > 90% of image.

Adequate macula and disc images when: 1) center of fovea (or disc) > 2DD from edge of image; 2) vessels clearly visible within 1DD of center of fovea (or disk), see Fig.8;

Inadequate macula and disc images when: failure to meet definition of Adequate unless diabetic retinopathy visible anywhere in the image.

A quality assessment of retinal images may be done in real time during the image registration. The main novelty of our study is idea to combine low level image processing procedures, content free quality assessment with medical grading of retinal images. We are testing now some algorithms for realization of the following scheme for retinal image quality assessment:

- 1) Is circular object in the image represent retina? If no, end.
- 2) Is the image color and contrast good enough or may be improved? If no, end.
- 3) Are vessel visibility good enough across of the image? If no, end.
- 4) Is macula or optic disc less than 2DD from the image center? If no, end.
- 5) If all conditions are met the studied image has good or adequate quality.

Conclusion

We presented our preliminary results of digital image processing oriented for retinal image quality analysis. Main problems in this way are lack of formal retinal quality definitions.

References

1. Gan, D. Diabetes atlas / D. Gan. – International Diabetes Federation, 2003.
2. Simple Anatomy of the Retina by Helga Kolb [Electronic resource]. – Mode of access : <http://webvision.med.utah.edu/book/part-i-foundations/simple-anatomy-of-the-retina/>. – Date of access : 14.02.2014.
3. Retinal image analysis: concepts, applications and potential / N. Patton [et al.] // Progress in retinal and eye research. – 2006. – Vol. 25, № 1. – P. 99–127.
4. Automated quality assessment of retinal fundus photos / J. Paulus [et al.] // International Journal of Computer Assisted Radiology and Surgery. – 2010. – Vol. 5, № 6. – P. 557–564.
5. Automated assessment of diabetic retinal image quality based on clarity and field definition / A.D. Fleming [et al.] // Investigative Ophthalmology and Visual Science. – 2006. – Vol. 47, № 3. – P. 1120–1125.
6. Elliptical local vessel density: a fast and robust quality metric for retinal images / L. Giancardo [et al.] // Proc. of 30th Annual Intern. Conf. Engineering in Medicine and Biology Society. – Vancouver, BC Canada, 2008. – P. 3534–3537.
7. Lalonde, M. Automatic visual quality assessment in optical fundus images / M. Lalonde, L. Gagnon, M. Boucher // Proc. of Vision Interface. – Ottawa, Ontario, Canada, 2001. – P. 259–264.
8. Lee, S.C. Automatic retinal image quality assessment and enhancement / S.C. Lee, Y. Wang // Proc. of Medical Imaging Conference. – USA, 1999. – Vol. 3661. – P. 1581–1590.
9. Image structure clustering for image quality verification of color retina images in diabetic retinopathy screening / M. Niemeijer [et al.] // Medical Image Analysis. – 2006. – Vol. 10, № 6. – P. 888–898.
10. Automated diagnosis and image understanding with object extraction, object classification, and inferencing in retinal images / Goldbaum M. [et al.] // Proc. of the Intern. Conf. on Image Processing. – Lausanne, Switzerland, 1996. – Vol. 3. – P. 695–698.

METHOD OF MOBILE-ROBOT CONTROL TO IMPLEMENT SEARCH MOTION

U. Sychyou

United Institute of Informatics Problems of the NAS of Belarus, Minsk
e-mail: vsychyov@robotics.by

In the article is discussed the problem of controlling a mobile robots collective for the dispersal on the space and re-transmission of information in order to increase the range of radio communication, to organize remote monitoring, and to control robot and mobile systems. The possible use of a chaotic oscillator is investigated to generate control signals for a mobile robot performing the search motion. Preprocessing method for data transmission via a radio channel has been developed. Computer simulation of proposed method search motion and field testing are carried out using mobile robot and robot arena for tests.

Introduction

Recently researchers have shown an increased interest in such an approach to the control of multirobot systems as swarm control. Swarm robots are considered to be called the mobile robots (MR) operating in groups under decentralized control with no direct data communication among them. Each robot can receive information on actions of the other robots only by measuring the state of the environment in which the group operates [1].

A distinctive feature of swarm control of MR, compared with other methods of group control, is the ability to achieve the goal, set before the group, by a large number of simple robots. Swarm control does not impose high demands on the quality of communication, reliability of hardware and quality of sensors. It is highly resistant to counteraction, as failure of a part of the group will not prevent the remaining robots to perform the task given.

The problem of group control of robots is formed similarly to the general problem of group control of a robots group [1]. Group control of robots originates in industrial robotics in which control systems of robots are captured by feedback, and robots operate in a determined environment. Determinacy of an environment means that the future state of an environment is completely defined by the current state and the action executed by a robot [2].

Environmental determinacy is reached due to limitation and closeness of a working zone in which there may not be objects with purpose and location unknown to a robot control system. At the same time, a robot's position in space is always exactly known. However usage of swarm robots is most favorable for solving tasks in a non-determined environment. Such tasks include inspecting a territory, relaying radio signals, etc. Besides, robots have to operate under conditions of incomplete and inaccurate information. E.g., for robots operating indoors it is often difficult with enough accuracy to determine their location and orientation in space without use of sophisticated navigation devices. Fig. 1 illustrates the operation of the obstacle sensor. In the one case (fig. 1, a), the sensor operates correctly, an error occurs in the second case (fig. 1, b). To sum up, it is possible to give the following classification of robots according to opportunities of positioning in space of a working zone:

1. Positioning in a working zone is impossible.
2. Positioning is possible relative to other robots.
3. Positioning is possible relative to a working zone.

Thus, application of swarm methods of control, in particular for control of microrobots or groups of simple robots belonging to the first category of the above classification, demands developing new control algorithms, the most important of which is the algorithm of search motion. The purpose of this work is to create an algorithm of search motion of a swarm robot.

The main requirement for this algorithm is its feasibility on robots equipped with only those sensors that are needed to solve a practical task.

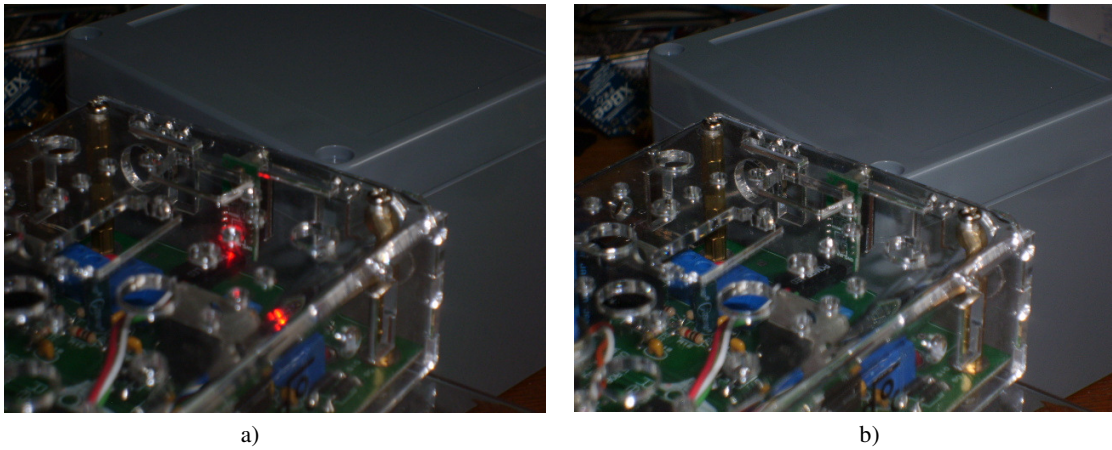


Fig. 1. Operation of the obstacle sensor: sensor operates correctly (a), an error occurs (b)

E.g., the algorithm must provide control of a robots group intended for relaying radio signals and not equipped with any sensors, except the received signal strength indicator (RSSI) and a front obstacle sensor.

1. Application of a nonlinear oscillator for control of mobile robot

The paper [3] describes a general scheme of search adaptive behavior with inertial switching between search tactics. As it was pointed out in this work, effective search motion can be realized by combination of moving for considerable distances, frequent random changes in motion directions, and persistence of switch between all types of motion.

The advantage of the described model of behavior is the possibility to implement search movements in conditions of incomplete and unreliable sensor information, which may be caused by disturbances, wrong operations of sensors or their dead zones.

The disadvantage of this method of search motion, applied to problems of robots control, consists in the necessity to program such a type of behavior by means of conditional transition functions, which may cause some difficulties in modification or control of coefficients of the operating program.

This disadvantage is not peculiar to the model described in [4]. It is based on the principle of controlling a mobile robot using a nonlinear oscillator, i.e. a system with dynamics described by differential equations of a second order [5]. The approach has a number of advantages in comparison with the reactive control method. Reactive control is widespread in robotics. Basing only on signals from sensors, it allows to realize robots control without using a model of the robot's operation environment. One of the most important benefits of the proposed approach consists in the opportunity to operate a nonlinear oscillator by means of a element which combines signals from a set of sensors. Thus, it becomes no longer required to logically program such actions as a robot's turn, stop, etc.

Besides, a nonlinear oscillator, unlike a pseudorandom sizes generator, allows to receive output signals of a various form – stationary, periodic and chaotic. These signals can be used for formation of various trajectories of MR motion.

In the paper it is offered to combine approaches [3] and [4] for realization of search motion of a mobile robot with use of controlling signals generated by a nonlinear oscillator. It will allow to implement a behavior model [3] with transitions between conditions caused by various operating modes of a nonlinear oscillator.

In work [4] a signal from a two-dimensional nonlinear oscillator serves for controlling a differential two-wheeled drive of a mobile robot.

In the current paper a nonlinear oscillator offered in work [4] is replaced with a chaotic signals generator (CSG) realizing the acquainted logistic map (a so-called logistic parabola), in detail described in [6, 7]:

$$y = a - y_{t-1}^2. \quad (1)$$

Fig. 2 illustrates the functional scheme of a mobile robot control system for realization of the search motion model which is described.

At the control system's input the vector of sensor data S is taken:

$$s = \{s_1, s_2, \dots, s_n\}, \quad (2)$$

where s_1, s_2, s_n denote data from each of sensors, n is the total number of sensors.

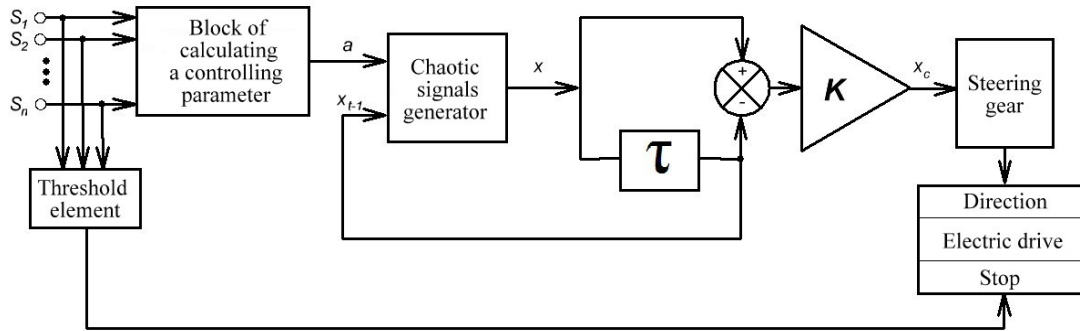


Fig. 2. The functional scheme of a mobile robot control system

The block of calculating a controlling parameter on the following formula calculates the parameter, controlling the CSG, on the basis of data from robot sensors:

$$a = w_1 s_1 + w_2 s_2 + \dots + w_n s_n, \quad (3)$$

where a is a controlling parameter, w_1, w_2, w_n are weight coefficients.

Thus, depending on value of the parameter a , the CSG can be in a stationary, periodic or chaotic mode [6]. Adjustment of weight coefficients allows to establish a correspondence between signals from sensors and a type of a desired trajectory of MR motion.

A threshold element allows to use a signal from any of sensors to stop a robot, e.g. in case of accident.

To convert output values of the CSG into control values of the MR steering gear, a scheme based on the element of signal delay in time τ , subtractor and proportional amplifier K , is used. This scheme is similar to a stabilization scheme of a nonlinear dynamical system with chaotic dynamics proposed by K. Piragas [7, 8]. At the output of the amplifier K the following signal is formed:

$$y_c = K(y_t - y_{t-1}). \quad (4)$$

The amplifier K allows to scale an output signal of the CSG for compatibility with various types of regulators of rotation speed of electric motors.

On the basis of the signal y_c power, brought to the left V_l and right V_r engines, is calculated according to the following rules:

$$N_l = -y_c, N_r = 225, \text{ in case } y_c < 0,$$

$$N_l = 225, N_r = y_c, \text{ in case } y_c > 0,$$

$$N_l = 225, N_r = 225, \text{ in case } y_c = 0.$$

The value of power N , brought to electric motors, can change in the range from 0 to 255. At $N = 255$ the left or right electric motor is supplied with the maximum power, and the respective crawler rotates forward at top speed, while at $N = 0$ – the same occurs in the opposite direction. The value $N = 128$ corresponds to zero power, and, therefore, a stop of electric motors. According to the (4), the signal y_c will be zero in case the CSG is in a stationary mode. In a periodic or chaotic regime the signal y_c will change respectively.

2. Computer modeling of a search motion algorithm

The object of computer modeling is a mobile robot equipped with a crawler differential drive [9] and controlled by means of the proposed algorithm. The purpose of modeling is to define influence of the CSG control parameter on the trajectory of a MR.

The trajectory of robot's motion, obtained as a result of computer modeling, is shown in fig. 3.

During modeling the control coefficient changed from $a = 0.2$ to $a = 1.9$.

The initial value was chosen $x = 1,1$.

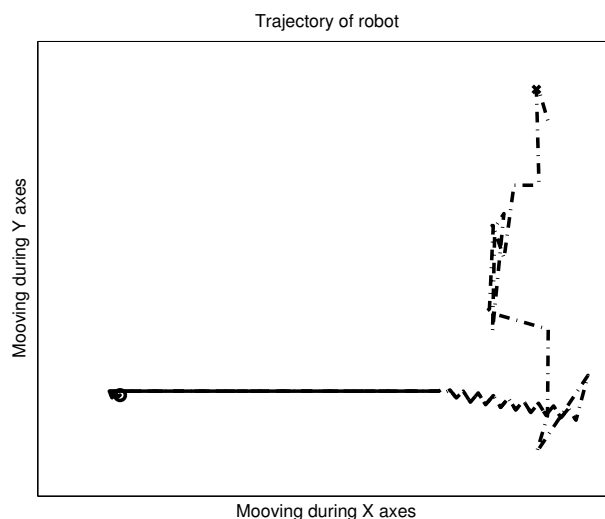


Fig. 3. The robot motion trajectory

The results of computer modeling confirm that a change of a MR motion trajectory can be reached by changing the operating coefficient. Besides, change of an operating mode of the CSG occurs gradually after change of the controlling parameter value. As a result, after response of sensors a robot needs some time for transition to another motion mode, and, therefore, short-term hindrances and false responses of sensors will not affect MR motion.

The proposed algorithm may find its practical application for solving the problem of radio signals retransmission using a group of swarm robots.

Conclusion

In this paper we proposed an algorithm of mobile robot search motion, based on the bionic model of adaptive searching behavior of V.A. Nepomnyashchikh [3], which was implemented on the basis of a chaotic signals generator. The proposed algorithm has important advantages, allowing, in particular, to build simple mobile robots equipped with a minimal number of sensors, and at the same time successfully solving search problems on unknown terrain.

References

1. Kaliaev, I. Models and algorithms of collective control in groups of robots / I. Kaliaev, A. Gaiduk, C. Kapustian. – Moscow : Phisimatlit, 2009. – 280 p. [In Russian]
2. Russell, S. Artificial Intelligence: A Modern Approach / S. Russell, P. Norvig; 2nd ed. – Prentice Hall PTR, 2002. – 1132 p.
3. Nepomnyashchikh, V.A Bionic model of adaptive searching behavior / V.A. Nepomnyashchikh, E.E. Popov, V.G. Red'ko // Journal of Computer and Systems Sciences International. – 2008. – № 1. – P. 85–93.
4. Clarck, M. Coupled Oscillator Control of Autonomous Mobile Robots / M. Clarck, T. Anderson, R. Skinner // Autonomous Robots. – Kluwer Academic Publishers, 2000. – P. 189–198.
5. Kovacic, I. Brennan The Duffing Equation : Nonlinear Oscillators and their Behaviour / I. Kovacic, M.J. Brennan. – John Wiley&Sons, 2011. – 386 p.
6. Moon, F. Chaotic Vibrations: An Introduction for Applied Scientists and Engineers / F. Moon. – John Wiley&Son, 2004. – 309 p.
7. Schöll, E. Schuster H.G. Handbook of Chaos Control / E Schöll, H.G. Schuster; 2nd ed. – Wiley-VCH Verlag GmbH&Co.KGaA, 2008. – 819 p.
8. Pyragas, K. Control of Chaos via an Unstable Delayed Feedback Controller / K. Pyragas. – Phys. Rev. Lett. 86, 2001. – P. 2265–2268.
9. Dudek, G. Computational Principles of Mobile Robotics / G. Dudek, M. Jenkin. – Cambridge University Press, 2000. – 280 p.

REGION GROWING SEGMENTATION OF CT-IMAGE BY THE ANALYSIS OF 3D LOCAL NEIGHBORHOOD

S. Trukhan², A. Nedzved¹, S. Ablameyko²

¹United Institute of Informatics Problems of the NAS of Belarus, Minsk;

²Belarusian State University, Minsk

e-mail: nedzveda@tut.by

New approach of matter segmentation in medical images is proposed. The main idea is rely on region growing method and additional constraints on inclusion of neighbor voxels. Additional constraints is based on analysis of 3D local neighborhood which is include only those voxels that satisfy the user specified thresholds on first and second derivatives. The paper will describe the implementation of widely used region growing algorithms in open-source.

Introduction

The growth of technologies favors creating much powerful medical equipment like Computed tomography (CT), Magnetic resonance tomography (MRT) and Positron emission tomography (PET). Current multislice CT scanners can configure up to 256 or even 512 detectors in an array. The resolution and quality of obtained images reaches remarkable values. Consider a standard chest CT image exam, which covers between 300 and 400 mm generating from 150 to 200 2-mm slices and up to 600 to 800 0.5-mm slices, depending on the slice thickness, or data sizes from 75 MB up to 400 MB. A whole body CT scan for screening can produce up to 2500 images or 1250 MB (1.25 GB) of data, with each image being 512×512×2 bytes [1]. With such a detailed data sets, the quality of image processing algorithms plays a significant part in clinical diagnostics but sometimes it is reached to the prejudice of performance.

There is no universal algorithm for segmentation of every medical image. Each imaging system has its own specific limitations. For example, in MR imaging (MRI) one has to take care of bias field noise (intensity in-homogeneities in the RF field). Of course, some methods are more general as compared to specialized algorithms and can be applied to a wider range of data [2].

One of the fastest semi-automatic segmentation methods is region-growing approach. The main idea is to grow region starting from seed pixel by comparing neighborhood pixel intensity with current. In the simplest case, a matter is in choosing of pixel, scanning of neighbors to find close values and merge it into new region. Common criterion for region homogeneity is based on estimation of maximal difference of current pixel's intensity and average intensity of new region. However, this clause will work if estimation of mean intensity is reliable only, so the size of region should not be small.

Among methods for region growing next are marked out: centroid connection (a priori information based on seed points), "merge-and-split" (growing of initially chosen homogeneous regions), "watershed" (based on gradient of intensity of image), deformable templates (based on template matching, which may change under function of internal energy). In addition, D.J. Withey and Z.J. Koles have provided a brief survey of three generations of medical image segmentation techniques [3].

1. Region growing in segmentation software

Region growing algorithms was implemented in many libraries and frameworks including open, commercial and academic sources. In some cases, it is better to use their highly tested and fast frameworks. The most popular and freely available medical imaging and processing frameworks are VTK and ITK. The main objective of VTK is data visualization but it still covers many image-processing algorithms (<http://vtk.org>). ITK provides developers with an extensive suite of software tools for image analysis and employs algorithms for registering and segmenting multidimensional data (<http://itk.org/>). Connected Threshold Image Filter, Neighborhood Connected Image Filter, Confidence Connected Image Filter are frequently used region growing ITK filters.

Connected Threshold Image Filter. This filter uses the flood fill iterator. Most of the algorithmic complexity of a region growing method comes from visiting neighboring pixels. The flood fill iterator assumes this responsibility and greatly simplifies the implementation of the region-growing algorithm. Thus, the algorithm is left to establish a criterion to decide whether a particular pixel should be included in the current region or not.

The criterion used by the Connected Threshold Image Filter is based on an interval of intensity values provided by the user (fig. 1). Values of lower and upper threshold should be provided. The region-growing algorithm includes those pixels whose intensities are inside the interval $I(X) \in [lower, upper]$.

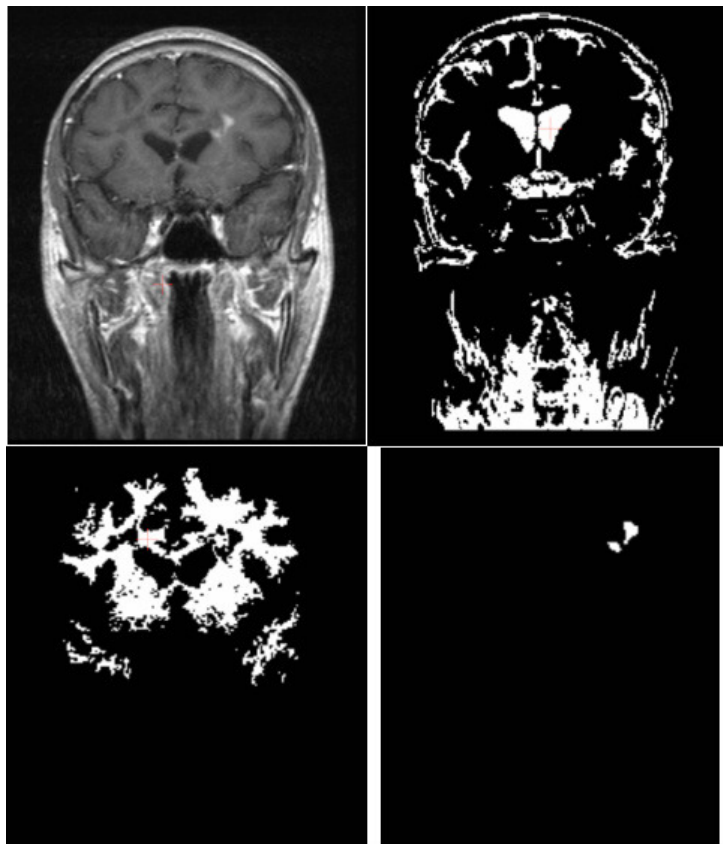


Fig. 11. Results of Connected Threshold Image Filter

Similar algorithm is implemented in VTK, `vtkImageThresholdConnectivity` class provides such functionality. Output results of ITK and VTK segmentation are identical, so in this case preferable to use VTK filter when the visualization part is made by VTK and not to spend memory and CPU time to converting pipeline from VTK to ITK and vice versa (table 1).

Table 1

Parameters used for segmenting some brain structures

Structure	Seed Index	Lower	Upper	Output Image (fig. 1)
White matter	(128, 208, 18)	633,00	826,20	Down left
Ventricle	(198, 146, 18)	154,00	440,20	Up right
White spot	(174, 210, 18)	973,00	6648,20	Down right

Neighborhood Connected Image Filter. This filter is a close variant of the Connected Threshold Image Filter. On one hand, the Connected Threshold Image Filter accepts a pixel in the region if its intensity is in the interval defined by two user-provided threshold values. The Neighborhood Connected Image Filter, on the other hand, will only accept a pixel if all its neighbors have intensities that fit in the interval. The size of the neighborhood to be considered around each pixel is defined by a user-provided integer radius.

The reason for considering the neighborhood intensities instead of only the current pixel intensity is that small structures are less likely to be accepted in the region. The operation of this filter is equivalent to applying the Connected Threshold Image Filter followed by mathematical morphology erosion using a structuring element of the same shape as the neighborhood provided to the Neighborhood Connected Image Filter.

Confidence Connected Image Filter. The criterion used by the Confidence Connected Image Filter is based on simple statistics of the current region. First, the algorithm computes the mean and standard deviation of intensity values for all the pixels currently included in the region. A user-provided factor is used to multiply the standard deviation and define a range around the mean. Neighbor pixels whose intensity values fall inside the range are accepted and included in the region. When no more neighbor pixels are found that satisfy the criterion, the algorithm is considered to have finished its first iteration. At that point, the mean and standard deviation of the intensity levels are recomputed using all the pixels currently included in the region. This mean and standard deviation defines a new intensity range that is used to visit current region neighbors and evaluate whether their intensity falls inside the range. This iterative process is repeated until no more pixels are added or the maximum number of iterations is reached.

The following equation illustrates the inclusion criterion used by this filter:
 $I(X) \in [m - f\sigma, m + f\sigma]$.

Where m and σ are the mean and standard deviation of the region intensities, f is a factor defined by the user, I is the image and X is the position of the particular neighbor pixel being considered for inclusion in the region.

2. Region growing segmentation based on the analysis of 3D local neighborhood

The main idea of introduced method rely on region growing method and constraints of inclusion of neighbor voxels. The analysis of 3D local neighborhood gives us two additional constraints.

The first stage is visiting neighborhood voxels starting from the seeds.

The second stage is to apply the constraints on neighboring voxels to the current. Neighbor voxel would be included to the current region if

1. $I(x, y) \in [lower, upper]$, where $I(x, y)$ – intensity of neighbor voxel, **lower and upper** – user specified thresholds.

2. $I'(x, y) \in [lower, upper]$, where $I'(x, y) = \left(\frac{\partial I}{\partial x}\right)^2 + \left(\frac{\partial I}{\partial y}\right)^2$ – the gradient magnitude of intensity of neighbor voxel, **lower and upper** – user specified thresholds.

3. $I''(x, y) \in [lower, upper]$, where $I''(x, y) = \left|\frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2}\right|$ – the Laplacian absolute value of intensity of neighbor voxel, **lower and upper** – user specified thresholds.

Algorithm is similar to Connected Threshold Image Filter with additional derivatives terms. It is remarkable that the computing of gradient would increase the computation time only by the constant value. Scheme of algorithm is on fig. 2.

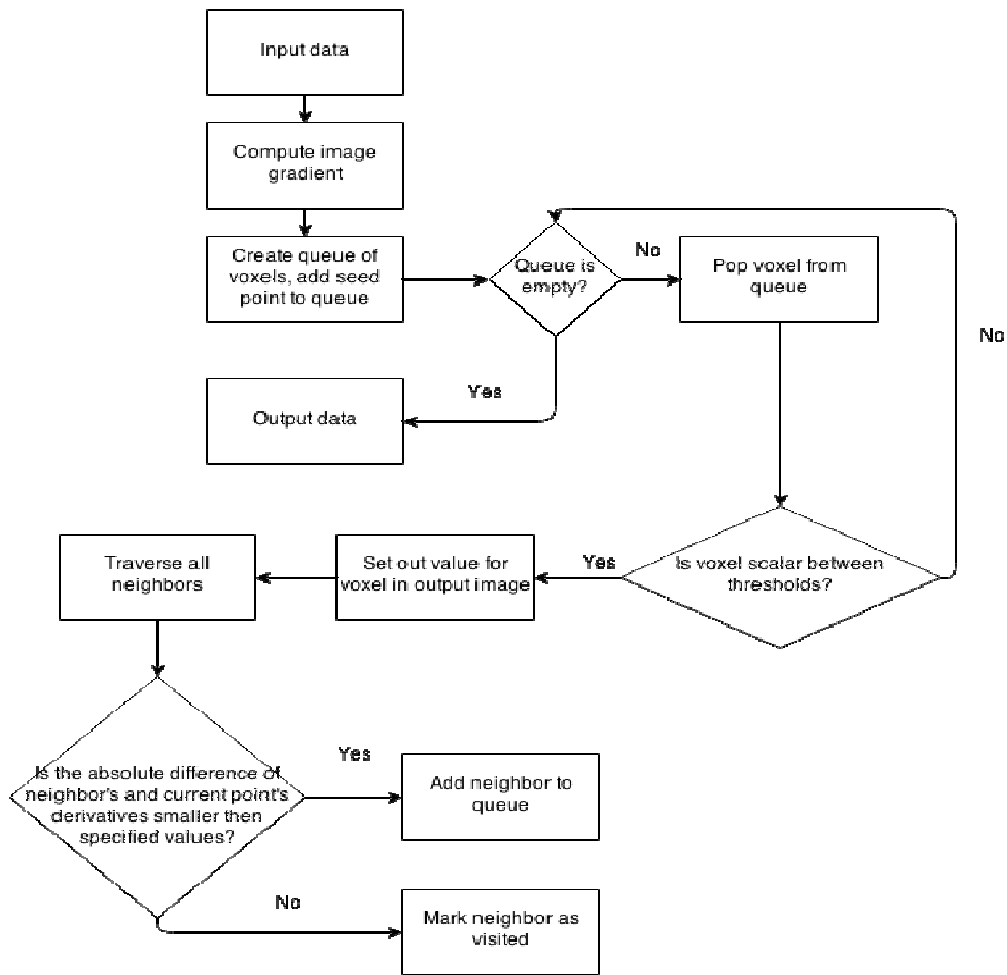


Fig. 12. Region growing segmentation based on the analysis of 3d local neighborhood

Conclusion

Region growing approach could be used with Energy functions, Bayesian functions, wavelets and fractals, and Neural net and may be performed in several ways starting from simple pixel-wise growing to models similar to active contours and “Snakes”. However, variety region growing methods either do not use a number of important local characteristics or work slowly. Therefore, new analysis should be introduced. One of new approaches is to analyze global and local information of image.

In medical images segmentation we often could know some image characteristics of region of interest. The first and second derivatives of each point could be computed using Sobel or Laplace filters. Analysis of these characteristics gives us more accuracy and quality of segmentation. The idea is not to add pixels which absolute difference between derivatives of current and neighbor pixel is larger than specified value (table 2).

Table 2

Proposed method						
Structure	Seed Index	Lower	Upper	First derivative	Second derivative	Output Image (fig. 3)
White matter	(128, 208, 18)	633,00	826,20	105,00	102,00	Down left
Ventricle	(198, 146, 18)	154,00	440,20	111,00	90,00	Up right
White spot	(174, 210, 18)	973,00	6648,20	164,00	180,00	Down right

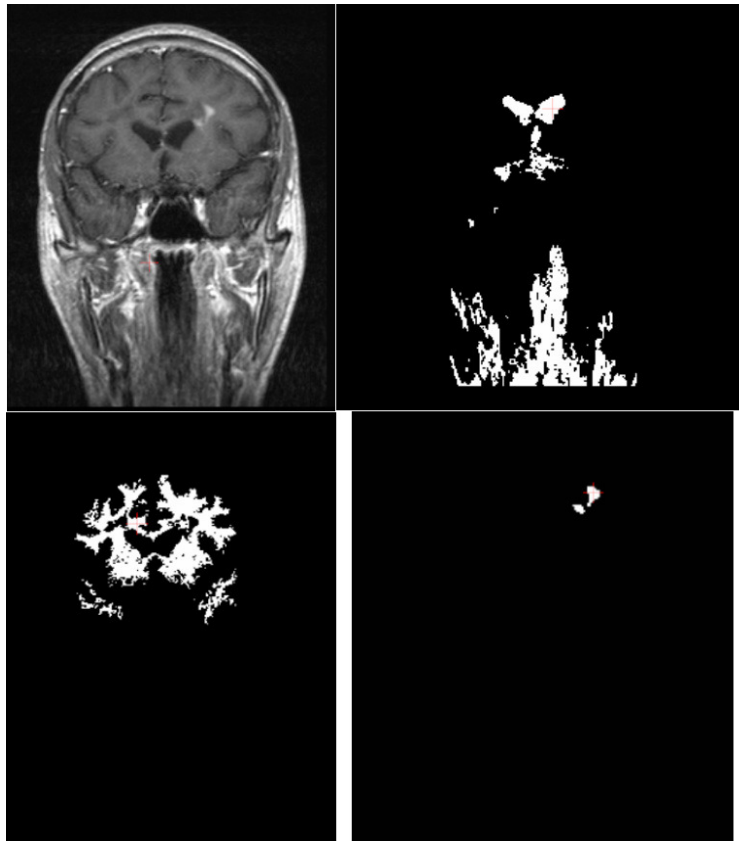


Fig. 13. Segmentation results for the region-growing algorithm with derivatives analysis

Comparing results of Connected Threshold Image Filter segmentation (fig. 1) and Region growing segmentation based on the analysis of 3D local neighborhood (fig. 3) we can admit that proposed method has higher accuracy.

References

1. Berezsky, O. Modern Trends in Biomedical Image Analysis System Design / O. Berezsky, G. Melnyk, Y. Batko // Biomedical Engineering, Trends in Electronics, Communications and Software. – 2011. – P. 24–44.
2. Casebased Medical Learning in Radiological Decision Making Using Contentbased Image Retrieval / P. Welter [et al.] // BMC Medical Informatics and Decision Making. – 2011. – Vol. 11. – P. 1472–6947.
3. Withey, D.J. Medical Image Segmentation: Methods and Software / D.J. Withey, Z.J. Koles // Noninvasive Functional Source Imaging of the Brain and Heart and the International Conference on Functional Biomedical Imaging, 2007. – China, Hangzhou, 2007. – P. 140–143.
4. Deserno, T.M. Fundamentals of Biomedical Image Processing / T.M. Deserno. – Berlin : Springer-Verlag, 2011. – 567 p.
5. Kaur, M. Medical Image Segmentation using Marker Controlled Watershed Transformation / M. Kaur, J. Jindal // Intern. J. of Computer Science & Technology. – 2011. – Vol. 2, no. 4. – P. 548–551.
6. Li, R. Medical Image Segmentation Based on Watershed Transformation and Rough Sets / R. Li // Bioinformatics and Biomedical Engineering (iCBBE) : 4th Intern. Conf. – NJ, Piscataway, 2010. – P. 1–5.
7. Morphogenesis-Based Deformable Models Application to 3D Medical Image Segmentation and Analysis / L. Ibanez [et al.] // Proceeding MICCAI '01 Proceedings of the 4th Intern. Conf. on Medical Image Computing and Computer-Assisted Intervention, 2001. – London : Springer-Verlag, 2001. – P. 1369–1370.

AN OUTLINE FOR AN APPROACH TO AUTOMATIC LABELING FOR INTERPRETATION OF HEURISTIC POSSIBILISTIC CLUSTERING RESULTS

D.A. Viattchenin¹, A. Damaratski¹, E. Nikolaenya¹, S. Shyrai²

¹United Institute of Informatics Problems of NAS of Belarus, Minsk

²Belarusian State University of Informatics and Radioelectronics, Minsk

e-mail: viattchenin@mail.ru

The paper deals in the preliminary way with the problem of clustering results interpretation for a case of application of heuristic algorithms of possibilistic clustering to data processing. The labeling problem and basic concepts the heuristic approach to possibilistic clustering are considered in brief. A labeling procedure is outlined. Experimental results are presented shortly and some preliminary conclusions are made.

Introduction

Clustering procedures belong to the algorithmic methods of data analysis. Various different approaches have been proposed for developing of classifiers by means of clustering, including:

- heuristic methods;
- optimization methods;
- hierarchical methods.

The first aim of clustering is to find structures contained within groups of data. These structures are usually classes to which objects from the data set are assigned. The result of the classification process is usually used as a classifier. Object, which as of yet have no known class assignments, are assigned to classes using the classifier constructed through clustering.

All clustering methods can be divided into three approaches, including:

- hard (or, in other word, crisp) methods;
- fuzzy methods;
- possibilistic methods.

Hard clustering methods assign each object to exactly one class, whereas in fuzzy and possibilistic clustering methods are assigned different degrees of membership to the different classes. In this way, a representativeness of each object for the single classes is determined during the analysis. So, the problem of clustering results interpretation is arises.

The main goal of this paper is a consideration of an approach to automatic labeling fuzzy clusters obtained from heuristic algorithms of possibilistic clustering. The contents of this paper is as follows: in the first section a labeling problem is described, in the second section basic concepts of the heuristic approach to possibilistic clustering are considered, in the third section a labeling procedure for clustering results interpretation is described, in the fourth a numerical example of application of the proposed algorithm to the Anderson's Iris data set are given. Some final remarks are also stated.

1. Fuzzy clustering results and labeling problem

The most widespread approach in fuzzy clustering is the optimization approach. Most optimization fuzzy clustering algorithms aim at minimizing an objective function that evaluates the partition of the data into a given number of fuzzy clusters. Objective function-based fuzzy clustering algorithms can in general be divided into two types: object versus relational.

The object data clustering methods can be applied if the objects are represented as points in some multidimensional space $I^{m_i}(X)$. In other words, the data which is composed of n objects and m_i attributes is denoted as $\hat{X}_{n \times m_i} = [\hat{x}_i^t]$, $i=1, K, n$, $t_1=1, K, m_i$ and the data are called sometimes the two-way data [1]. Let $X = \{x_1, \dots, x_n\}$ is the set of objects. So, the two-way data matrix can be represented as follows:

$$\hat{X}_{n \times m_i} = \begin{pmatrix} \hat{x}_1^1 & \hat{x}_1^2 & \dots & \hat{x}_1^{m_i} \\ \hat{x}_2^1 & \hat{x}_2^2 & \dots & \hat{x}_2^{m_i} \\ \dots & \dots & \dots & \dots \\ \hat{x}_n^1 & \hat{x}_n^2 & \dots & \hat{x}_n^{m_i} \end{pmatrix}. \quad (1)$$

So, the two-way data matrix can be represented as $\hat{X} = (\hat{x}^1, K, \hat{x}^{m_i})$ using n -dimensional column vectors \hat{x}^i , $i = 1, K, m_i$, composed of the elements of the t_i -th column of \hat{X} .

The traditional optimization methods of fuzzy clustering are based on the concept of fuzzy c -partition [2]. The initial set $X = \{x_1, \dots, x_n\}$ of n objects represented by the matrix of similarity coefficients, the matrix of dissimilarity coefficients or the matrix of object attributes, should be divided into c fuzzy clusters. Namely, the grade u_{li} , $1 \leq l \leq c$, $1 \leq i \leq n$ to which an object x_i belongs to the fuzzy cluster A^l should be determined. For each object x_i , $i = 1, K, n$ the grades of membership should satisfy the conditions of a fuzzy c -partition:

$$\sum_{l=1}^c u_{li} = 1, \quad 1 \leq i \leq n, \quad 0 \leq u_{li} \leq 1, \quad 1 \leq l \leq c. \quad (2)$$

In other words, the family of fuzzy sets $P(X) = \{A^l \mid l = \overline{1, c}, c \leq n\}$ is the fuzzy c -partition of the initial set of objects $X = \{x_1, \dots, x_n\}$ if condition (2) is met. Fuzzy c -partition $P(X)$ may be described with the aid of a partition matrix $P_{c \times n} = [u_{li}]$, $l = 1, K, c$, $i = 1, K, n$. The set of all fuzzy c -partitions will be denoted by Π . So, the fuzzy problem formulation in cluster analysis can be defined as the optimization task $Q \rightarrow \underset{P(X) \in \Pi}{extr}$ under the constraints (2), where Q is a fuzzy objective function.

The best known optimization approach to fuzzy clustering is the method of fuzzy c -means [3]. The FCM-algorithm is based on an iterative optimization of the fuzzy objective function, which takes the form:

$$Q_{FCM}(P, T) = \sum_{l=1}^c \sum_{i=1}^n u_{li}^\gamma \|x_i - \tau^l\|^2, \quad (3)$$

where u_{li} , $l = 1, K, c$, $i = 1, K, n$ is the membership degree, x_i , $i \in \{1, K, n\}$ is the data point, $T = \{\tau^1, K, \tau^c\}$ is the set fuzzy clusters prototypes, and $\gamma > 1$ is the weighting exponent.

The purpose of the classification task is to obtain the solutions $P(X)$ and τ^1, K, τ^c which minimize equation (3). Some other similar objective function-based fuzzy clustering algorithms, such as the GK-algorithm, are considered in [3–5] in detail.

In order to be applying the found cluster prototype as classifiers, they need to be given reasonable names. One can then use these names as column titles of the membership matrix when using the recall function of the FCM-algorithm. This helps in the interpretation of the results.

The process of assigning class names to cluster prototypes is called labeling. A labeling method for the fuzzy c -means method is to inspect the cluster prototypes and their respective membership values of the various attributes, and to assign a label manually.

However, usually, it is already known when training a classifier which objects belong to which classes. This information can be taken into account to use so as automatic the fuzzy c -means labeling process. The corresponding labeling procedure is described in [6] in detail. The principal idea of the procedure is to present a sample of objects to the FCM-classifier whose class membership are known in the hard form of 0 or 1 values and have also been calculated by the procedure. By means of the given cluster membership values for each fuzzy cluster prototype, the fuzzy cluster prototypes can be associated with their respective classes.

On the other hand, all objective function-based fuzzy clustering algorithms are iterative procedures and the initial fuzzy c -partition $P(X)$ is initialized randomly. So, coordinates of fuzzy clusters prototypes and values of membership functions will be different in each experiment for the same data set, because the result of classification is sensitive to initialization. Moreover, major objective function-based fuzzy clustering algorithms are need for using some validity measures [2] for determining the most “plausible” number c of fuzzy clusters in the sought fuzzy c -partition $P(X)$. So, a problem of rapid automatic labeling is arises.

2. Basic concepts of the heuristic approach to possibilistic clustering

Let us remind the basic concepts of the heuristic method of possibilistic clustering [7]. Let $X = \{x_1, \dots, x_n\}$ be the initial set of elements and $T: X \times X \rightarrow [0,1]$ some fuzzy tolerance on X with $\mu_T(x_i, x_j) \in [0,1], \forall x_i, x_j \in X$ being its membership function. Let α be the α -level value of the fuzzy tolerance T , $\alpha \in (0,1]$. Columns or rows of the fuzzy tolerance matrix are fuzzy sets $\{A^1, \dots, A^n\}$ on X . Let $A^l, l \in \{1, K, n\}$ be a fuzzy set on X with $\mu_{A^l}(x_i) \in [0,1], \forall x_i \in X$ being its membership function. The α -level fuzzy set $A_{(\alpha)}^l = \{(x_i, \mu_{A^l}(x_i)) \mid \mu_{A^l}(x_i) \geq \alpha, x_i \in X\}$ is fuzzy α -cluster. So, $A_{(\alpha)}^l \subseteq A^l, \alpha \in (0,1], A^l \in \{A^1, K, A^n\}$ and $\mu_{A^l}(x_i)$ is the membership degree of the element $x_i \in X$ for some fuzzy α -cluster $A_{(\alpha)}^l, \alpha \in (0,1], l \in \{1, K, n\}$. The membership degree will be denoted μ_{li} in further considerations. The membership degree of the element $x_i \in X$ for some fuzzy α -cluster $A_{(\alpha)}^l, \alpha \in (0,1], l \in \{1, K, n\}$ can be defined as a

$$\mu_{li} = \begin{cases} \mu_{A^l}(x_i), & x_i \in A_{(\alpha)}^l \\ 0, & \text{otherwise} \end{cases}, \quad (4)$$

where the α -level $A_{(\alpha)}^l = \{x_i \in X \mid \mu_{A^l}(x_i) \geq \alpha\}, \alpha \in (0,1]$ of a fuzzy set A^l is the support of the fuzzy α -cluster $A_{(\alpha)}^l$ and value of α is the tolerance threshold of fuzzy α -cluster elements.

Let $\{A_{(\alpha)}^1, \dots, A_{(\alpha)}^n\}$ be the family of fuzzy α -clusters for some α . The point $\tau_e^l \in A_{(\alpha)}^l$, for which

$$\tau_e^l = \arg \max_{x_i} \mu_{li}, \forall x_i \in A_{(\alpha)}^l, \quad (5)$$

is called a typical point of the fuzzy α -cluster $A_{(\alpha)}^l, \alpha \in (0,1], l \in [1, n]$. A set $K(A_{(\alpha)}^l) = \{\tau_{|l}^l, K, \tau_{|l}^l\}$ of typical points of the fuzzy cluster $A_{(\alpha)}^l$ is a kernel of the fuzzy cluster and $\text{card}(K(A_{(\alpha)}^l)) = |l|$ is a cardinality of the kernel. If the fuzzy cluster have an unique typical point, then $|l| = 1$.

Let $R_z^\alpha(X) = \{A_{(\alpha)}^l \mid l = \overline{1, c}, 2 \leq c \leq n\}$ be a family of fuzzy α -clusters for some value of tolerance threshold α , which are generated by a fuzzy tolerance T on the initial set of elements $X = \{x_1, \dots, x_n\}$. If condition

$$\sum_{l=1}^c \mu_{li} > 0, \forall x_i \in X, \quad (6)$$

is met for all $A_{(\alpha)}^l, l = \overline{1, c}, c \leq n$, then the family is the allotment of elements of the set $X = \{x_1, \dots, x_n\}$ among fuzzy α -clusters $\{A_{(\alpha)}^l, l = \overline{1, c}, 2 \leq c \leq n\}$ for some value of the tolerance threshold α . It should be noted that several allotments $R_z^\alpha(X)$ can exist for some tolerance threshold α . That is why symbol z is the index of an allotment.

Obviously, the definition of the allotment among fuzzy clusters (6) is similar to the definition of the possibilistic partition [8]. So, the allotment among fuzzy clusters can be considered as the possibilistic partition and fuzzy clusters in the sense of (4) are elements of the possibilistic partition.

Thus, the problem of cluster analysis can be defined as the problem of discovering the unique allotment $R_c^*(X)$, resulting from the classification process and detection of fixed or unknown number c of fuzzy α -clusters can be considered as the aim of classification.

Direct heuristic algorithms of possibilistic clustering can be divided into two types: relational versus prototype-based. A fuzzy tolerance relation matrix is a matrix of the initial data for the direct heuristic relational algorithms of possibilistic clustering and a matrix of attributes is a matrix for the prototype-based algorithms. These algorithms are described in [7] in detail. So, the matrix of memberships $R_c^*(X) = [\mu_{li}]$, the value α of the tolerance threshold and the set of kernels $\{K(A_{(\alpha)}^1), K, K(A_{(\alpha)}^c)\}$ are results of classification. The results will be constant in each experiment for the same data set, because the sought clustering structure of the set of objects is based directly on the formal

definition of fuzzy cluster and the possibilistic memberships are determined directly from the values of the pairwise similarity of objects.

3. A labeling procedure

The results obtained from heuristic algorithms of possibilistic clustering are stable. The set of kernels $\{K(A_{(\alpha)}^1), K, K(A_{(\alpha)}^c)\}$ and the set of labels $\{label 1, K, label c\}$ are inputs for the procedure. We assume that a condition $card(K(A_{(\alpha)}^l))=1$ is met for each kernel $K(A_{(\alpha)}^l), l = \overline{1, c}$. In other words, the set of typical points $\{\tau^1, K, \tau^c\}$ is given. So, the labeling procedure can be described as follows:

Perform the following operations for each typical point $\tau^l, l = \overline{1, c}$ and each label $label m, m = \overline{1, c}$:

1. Let $l := 1$ and $m := 1$;
2. Check the following condition:
 - 2.1 **if** τ^l corresponds to $label m$ **then** the label $label m$ is label for the typical point τ^l and go to step 3 **else** $m := m + 1$ and go to step 2;
3. Check the following condition:
 - 3.1 **if** the typical point τ^l is labeled **then** $l = l + 1$ and go to step 2 **else** go to step 4;
4. Check the following condition:
 - 4.1 **if** all typical points $\tau^l, l = \overline{1, c}$ are labeled **then** stop.

So, the proposed procedure is fast and very simply. Finally, a label can be assigned to all objects, which correspond to the cluster.

4. An illustrative example

The Anderson's Iris database [9] is the most known database to be found in the pattern recognition literature. The data set represents different categories of Iris plants having four attribute values. The four attribute values represent the sepal length, sepal width, petal length and petal width measured for 150 irises. It has three classes Setosa, Versicolor and Virginica, with 50 samples per class. An example of the records in the database is presented in table 1.

Table 1

An example of the records in the Iris database

Numbers of objects	Attributes				Labels of classes
	Sepal length	Sepal width	Petal length	Petal width	
...
18	5.1	3.3	1.7	0.5	Setosa
...

By executing the D-AFC(c)-algorithm [7] for $c = 3$ using the normalized Euclidean distance, we obtain that the typical point of the first class τ^1 is the object x_{23} , the typical point of the second class τ^2 is the object x_{95} , and the typical point of the third class τ^3 is the object x_{98} . The set of labels is $\{Setosa, Versicolor, Virginica\}$. The result of application of the proposed labeling procedure to clustering results is summarized in table 2.

Table 2

Typical points of fuzzy clusters

Numbers of clusters	Numbers of objects	Labels of classes
1	23	Virginica
2	95	Setosa
3	98	Versicolor

So, the interpreted result of the classification process is summarized in table 3, where misclassified objects are bolded.

Final results of the Iris data set processing

Numbers of classes	Labels of classes	Numbers of objects
1	Virginica	2, 4, 7, 9 , 13, 15, 16, 17, 20, 21, 23, 24, 27, 32, 34, 35, 39, 41, 45, 46, 49, 52, 57, 62, 73, 74, 75, 77, 80, 81, 82, 83, 89, 102, 104, 108, 110, 111, 122, 123, 126, 127, 131, 132, 137, 146, 147 , 148
2	Setosa	1, 6, 10, 18, 26, 31, 36, 37, 40, 42, 44, 47, 50, 51, 53, 54, 55, 58, 59, 60, 63, 64, 67, 68, 71, 72, 78, 79, 87, 88, 91, 95, 96, 100, 101, 106, 107, 112, 115, 124, 125, 134, 135, 136, 138, 139, 143, 144, 145, 149
3	Versicolor	3, 5 , 8, 11, 12, 14, 19, 22, 25 , 28, 29, 30, 33, 38, 43, 48, 56 , 61, 65, 66, 69, 70, 76, 84, 85, 86, 90 , 92, 93, 94, 97, 98, 99, 103, 105, 109, 113, 114, 116, 117, 118, 119, 120, 121, 128, 129, 130, 133, 140, 141, 142, 147, 150

Conclusions

The fast labeling procedure for interpretation of results of the data processing by using heuristic possibilistic clustering results is proposed. The procedure can be used for learning fuzzy inference systems, which are generated by using heuristic possibilistic clustering [7].

References

1. Sato-Ilic, M. Innovations in Fuzzy Clustering / M. Sato-Ilic, L.C. Jain. – Heidelberg : Springer, 2006. – 152 p.
2. Fuzzy Models and Algorithms for Pattern Recognition and Image Processing / J.C. Bezdek [et al.]. – N.Y. : Springer Science, 2005. – 776 p.
3. Bezdek, J.C. Pattern Recognition with Fuzzy Objective Function Algorithms / J.C. Bezdek. – N.Y. : Plenum Press, 1981. – 272 p.
4. Miyamoto, S. Algorithms for Fuzzy Clustering: Methods in C-Means Clustering with Applications / S. Miyamoto, H. Ichihashi, K. Honda. – Heidelberg : Springer, 2008. – 247 p.
5. Fuzzy Cluster Analysis: Methods for Classification, Data Analysis and Image Recognition / F. Höppner [et al.]. – Chichester : John Wiley & Sons, 1999. – 289 p.
6. DataEngine: Tutorials and Theory. – Aachen : MIT GmbH, 1999. – 255 p.
7. Viattchenin, D.A. A Heuristic Approach to Possibilistic Clustering: Algorithms and Applications / D.A. Viattchenin. – Heidelberg : Springer, 2013. – 227 p.
8. Krishnapuram, R. A Possibilistic Approach to Clustering / R. Krishnapuram, J.M. Keller // IEEE Transactions on Fuzzy Systems. – 1993. – Vol. 1 (1). – P. 98–110.
9. Anderson, E. The Irises of the Gaspé Peninsula / E. Anderson // Bulletin of the American Iris Society. – 1935. – Vol. 59 (1). – P. 2–5.

NEW PARAMETERS FOR RELATIONAL HEURISTIC ALGORITHMS OF POSSIBILISTIC CLUSTERING

D.A. Viattchenin¹, A. Damaratski¹, A. Yaroma²

¹United Institute of Informatics Problems of NAS of Belarus, Minsk

²Belarusian State University of Informatics and Radioelectronics, Minsk

e-mail: viattchenin@mail.ru

The theoretical note introduces new parameters for relational heuristic algorithms of possibilistic clustering. These parameters can be used for elaboration of new for relational heuristic algorithms of possibilistic clustering. Classification problems are formulated and some preliminary conclusions are made.

Introduction

Determining a partition of given sample data is an important part in data analysis tasks. Clustering methods use a mathematical model based on similarity measure to determine a suitable partition of the data set. In fuzzy clustering the data is not only partitioned in a number of clusters, but each object is assigned a degree of membership for each cluster.

Heuristic methods, hierarchical methods and objective function-based methods are main approaches in fuzzy clustering. In objective function-based clustering the mathematical model is stated in form of an objective function that evaluates the partition of data with respect to the membership degrees and the underlying similarity or dissimilarity measure. Different assumptions and constraints lead to a variety of basic clustering concepts. If the objective function is differentiable, necessary conditions for the membership degrees and other cluster parameters used in the distance or similarity measure can be derived in order to optimize the objective function. The resulting equations are then alternatively applied in an algorithm to determine the data fuzzy partition [1].

A possibilistic approach to clustering was proposed by Krishnapuram and Keller [2] and the approach can be considered as a special case of fuzzy approach to clustering because all methods of possibilistic clustering are objective function-based methods. On the other hand, constraints in the possibilistic approach to clustering are less strong than constraints in the fuzzy objective function-based approach to clustering and values of the membership function of a possibilistic partition can be considered as typicality degrees. So, the possibilistic approach to clustering is more general and flexible approach to clustering than the fuzzy approach.

Objective function-based approach in fuzzy clustering is most common and widespread approach. However, heuristic algorithms of fuzzy clustering display low level of complexity and high level of essential clarity. Some heuristic clustering algorithms are based on a definition of the cluster concept and the aim of these algorithms is cluster detection conform to a given definition. Such algorithms are called algorithms of direct classification or direct clustering algorithms. A heuristic approach to possibilistic clustering is proposed in [3].

The main purpose of this short note is a proposition of new parameters for heuristic algorithms of possibilistic clustering. For the purpose, the heuristic approach to possibilistic clustering is described in brief, basic definitions of the approach are considered, new parameters of classification are introduced and classification tasks are formulated, and preliminary conclusions are stated.

1. A heuristic approach to possibilistic clustering

The essence of the heuristic approach to possibilistic clustering is that the sought clustering structure of the set of observations is formed based directly on the formal definition of fuzzy cluster and possibilistic memberships are determined also directly from the values of the pairwise similarity of observations. A concept of the allotment among fuzzy clusters is basic concept of the approach and the allotment among fuzzy clusters is a special case of the possibilistic partition.

Direct heuristic algorithms of possibilistic clustering can be divided into two types: relational versus prototype-based. A fuzzy tolerance relation matrix is a matrix of the initial data for the direct heuristic relational algorithms of possibilistic clustering and a matrix of attributes is a matrix of the initial data for the prototype-based algorithms. In particular, the group of direct relational heuristic algorithms of possibilistic clustering includes

- D-AFC(c)-algorithm: using the construction of the allotment among given number c of partially separate fuzzy clusters;
- D-PAFC-algorithm: using the construction of the principal allotment among an unknown minimal number of at least c fully separate fuzzy clusters;
- D-AFC-PS(c)-algorithm: using the partially supervised construction of the allotment among given number c of partially separate fuzzy clusters.

On the other hand, the family of direct prototype-based heuristic algorithms of possibilistic clustering includes

- D-AFC-TC-algorithm: using the construction of the allotment among an unknown number c of fully separate fuzzy clusters;
- D-PAFC-TC-algorithm: using the construction of the principal allotment among an unknown minimal number of at least c fully separate fuzzy clusters;
- D-AFC-TC(α)-algorithm: using the construction of the allotment among an unknown number c of fully separate fuzzy clusters with respect to the minimal value α of the tolerance threshold.

It should be noted, that these prototype-based heuristic algorithms of possibilistic clustering are based on the transitive closure of initial fuzzy tolerance. New direct prototype-based heuristic algorithms of possibilistic clustering were proposed in [4] and the family of algorithms is based on the transitive approximation of fuzzy tolerance [5].

2. Basic concepts of the heuristic approach to possibilistic clustering

Basic concepts of the heuristic method of possibilistic clustering [3] must be reminded before introducing new parameters of classification.

Let $X = \{x_1, \dots, x_n\}$ be the initial set of objects. Let T be a fuzzy tolerance on X and α be α -level value of T , $\alpha \in (0,1]$. Columns or lines of the fuzzy tolerance matrix are fuzzy sets $\{A^1, \dots, A^n\}$. Let $\{A^1, \dots, A^n\}$ be fuzzy sets on X , which are generated by a fuzzy tolerance T . The α -level fuzzy set $A_{(\alpha)}^l = \{(x_i, \mu_{A^l}(x_i)) \mid \mu_{A^l}(x_i) \geq \alpha\}$, $l \in [1, n]$ is fuzzy α -cluster or, simply, fuzzy cluster. So $A_{(\alpha)}^l \subseteq A^l$, $\alpha \in (0,1]$, $A^l \in \{A^1, \dots, A^n\}$ and μ_{A^l} is the membership degree of the element $x_i \in X$ for some fuzzy cluster $A_{(\alpha)}^l$, $\alpha \in (0,1]$, $l \in [1, n]$. Value of α is the tolerance threshold of fuzzy clusters elements.

The membership degree of the element $x_i \in X$ for some fuzzy cluster $A_{(\alpha)}^l$, $\alpha \in (0,1]$, $l \in [1, n]$ can be defined as a

$$\mu_{A_{(\alpha)}^l} = \begin{cases} \mu_{A^l}(x_i), & x_i \in A_{(\alpha)}^l \\ 0, & \text{otherwise} \end{cases}, \quad (1)$$

where an α -level $A_{(\alpha)}^l = \{x_i \in X \mid \mu_{A^l}(x_i) \geq \alpha\}$, $\alpha \in (0,1]$ of a fuzzy set A^l is the support of the fuzzy cluster $A_{(\alpha)}^l$. So, condition $A_{(\alpha)}^l = \text{Supp}(A_{(\alpha)}^l)$ is met for each fuzzy cluster $A_{(\alpha)}^l$, $\alpha \in (0,1]$, $l \in [1, n]$. Membership degree can be interpreted as a degree of typicality of an element to a fuzzy cluster.

Let T is a fuzzy tolerance on X , where X is the set of objects, and $\{A_{(\alpha)}^1, \dots, A_{(\alpha)}^n\}$ is the family of fuzzy clusters for some $\alpha \in (0,1]$. The point $\tau_e^l \in A_{(\alpha)}^l$, for which

$$\tau_e^l = \arg \max_{x_i} \mu_{A_{(\alpha)}^l}, \quad \forall x_i \in A_{(\alpha)}^l, \quad (2)$$

is called a typical point of the fuzzy cluster $A_{(\alpha)}^l$, $\alpha \in (0,1]$, $l \in [1, n]$. A fuzzy cluster $A_{(\alpha)}^l$ can have several typical points. That is why symbol e is the index of the typical point.

Let $R_{c(z)}^\alpha(X) = \{A_{(\alpha)}^l \mid l = \overline{1, c}, 2 \leq c \leq n, \alpha \in (0, 1]\}$ be a family of fuzzy clusters for some value of tolerance threshold α , $\alpha \in (0, 1]$, which are generated by some fuzzy tolerance T on the initial set of elements $X = \{x_1, \dots, x_n\}$. If a condition

$$\sum_{i=1}^c \mu_{ii} > 0, \quad \forall x_i \in X \quad (3)$$

is met for all fuzzy clusters $A_{(\alpha)}^l \in R_{c(z)}^\alpha(X)$, $l = \overline{1, c}$, $c \leq n$, then the family is the allotment of elements of the set $X = \{x_1, \dots, x_n\}$ among fuzzy clusters $\{A_{(\alpha)}^l, l = \overline{1, c}, 2 \leq c \leq n\}$ for some value of the tolerance threshold α . It should be noted that several allotments $R_{c(z)}^\alpha(X)$ can exist for some tolerance threshold α . That is why symbol z is the index of an allotment.

Allotment $R_l^\alpha(X) = \{A_{(\alpha)}^l \mid l = \overline{1, n}, \alpha \in (0, 1]\}$ of the set of objects among n fuzzy clusters for some tolerance threshold $\alpha \in (0, 1]$ is the initial allotment of the set $X = \{x_1, \dots, x_n\}$. In other words, if initial data are represented by a matrix of some fuzzy T then lines or columns of the matrix are fuzzy sets $A^l \subseteq X$, $l = \overline{1, n}$ and α -level fuzzy sets $A_{(\alpha)}^l$, $l = \overline{1, c}$, $\alpha \in (0, 1]$ are fuzzy clusters. These fuzzy clusters constitute an initial allotment for some tolerance threshold α and they can be considered as clustering components.

If some allotment $R_{c(z)}^\alpha(X) = \{A_{(\alpha)}^l \mid l = \overline{1, c}, c \leq n, \alpha \in (0, 1]\}$ corresponds to the formulation of a concrete problem, then this allotment is an adequate allotment. In particular, if a condition

$$\sum_{l=1}^c \text{card}(A_{(\alpha)}^l) \geq \text{card}(X), \quad \forall A_{(\alpha)}^l \in R_{c(z)}^\alpha(X), \quad \alpha \in (0, 1], \quad \text{card}(R_{c(z)}^\alpha(X)) = c, \quad (4)$$

and a condition

$$\text{card}(A_{(\alpha)}^l \cap A_{(\alpha)}^m) \leq w, \quad \forall A_{(\alpha)}^l, A_{(\alpha)}^m, \quad l \neq m, \quad \alpha \in (0, 1], \quad (5)$$

are met for all fuzzy clusters $A_{(\alpha)}^l$, $l = \overline{1, c}$ of some allotment $R_{c(z)}^\alpha(X) = \{A_{(\alpha)}^l \mid l = \overline{1, c}, c \leq n\}$ for a value $\alpha \in (0, 1]$, then the allotment is the allotment among partially separate fuzzy clusters.

Allotment $R_p^\alpha(X) = \{A_{(\alpha)}^l \mid l = \overline{1, c}\}$ of the set of objects among the minimal number c , $2 \leq c \leq n$ of fully separate fuzzy clusters for some tolerance threshold $\alpha \in (0, 1]$ is the principal allotment of the set $X = \{x_1, \dots, x_n\}$.

Several adequate allotments can exist. Thus, the problem consists in the selection of the unique adequate allotment $R_c^*(X)$ from the set B of adequate allotments, $B = \{R_{c(z)}^\alpha(X)\}$, which is the class of possible solutions of the concrete classification problem. The selection of the unique adequate allotment $R_c^*(X)$ from the set $B = \{R_{c(z)}^\alpha(X)\}$ of adequate allotments must be made on the basis of evaluation of allotments. The criterion

$$F(R_{c(z)}^\alpha(X), \alpha) = \sum_{l=1}^c \frac{1}{n_l} \sum_{i=1}^{n_l} \mu_{ii} - \alpha \cdot c, \quad (6)$$

where c is the number of fuzzy clusters in the allotment $R_{c(z)}^\alpha(X)$ and $n_l = \text{card}(A_{(\alpha)}^l)$, $A_{(\alpha)}^l \in R_{c(z)}^\alpha(X)$ is the number of elements in the support of the fuzzy cluster $A_{(\alpha)}^l$, can be used for evaluation of allotments. Maximum of criterion (6) corresponds to the best allotment of objects among c fuzzy clusters. So, the classification problem can be characterized formally as determination of the solution $R_c^*(X)$ satisfying

$$R_c^*(X) = \arg \max_{R_{c(z)}^\alpha(X) \in B} F(R_{c(z)}^\alpha(X), \alpha). \quad (7)$$

The problem of cluster analysis can be defined in general as the problem of discovering the unique allotment $R_c^*(X)$, resulting from the classification process.

3. Parameters of classification

The number of classes c in an allotment can be unknown a priori. So, a problem of detecting the unique allotment among unknown number c of fuzzy clusters may arise. On the other hand, though, some other information about either classification nature or properties of classes can be known a priori. Let us consider in detail parameters of classification.

Firstly, a cluster is a non-empty set of elements. So, an analyst can determine the maximal number u of elements in a fuzzy cluster. If $1 \leq u < n$ is a maximal number of elements in a fuzzy cluster, then $1 \leq n_l \leq u$, $\forall l = \overline{1, c}$, where $n_l = \text{card}(A_\alpha^l)$, $A_\alpha^l = \text{Supp}(A_{(\alpha)}^l)$ for each fuzzy cluster $A_{(\alpha)}^l$, $l = \overline{1, c}$, $\alpha \in (0, 1]$. So, parameter u can be considered as the parameter that controls cluster sizes. This natural idea was developed for the FCM-algorithm by Miyamoto, Ichihashi and Honda [6].

Secondly, the minimal value of the tolerance threshold can be determined a priori, so that allotments $R_{c(z)}^\alpha(X) = \{A_{(\alpha)}^l \mid l = \overline{1, c}, c \leq n, \alpha \in (0, 1]\}$ are constructed for every α , $\alpha \in [\alpha, 1]$, where α is the tolerance threshold, estimated by the analyst.

Thirdly, the maximal number \hat{w} , $0 \leq \hat{w} < n$ of elements in the intersection area of fuzzy clusters can be considered as a parameter of a direct algorithm of possibilistic clustering. So, the condition (5) can be rewritten as follows:

$$\text{card}(A_\alpha^l \cap A_\alpha^m) \leq \hat{w}, \forall A_{(\alpha)}^l, A_{(\alpha)}^m, l \neq m, \alpha \in (0, 1]. \quad (8)$$

Thus, \hat{w} can be considered as the parameter that controls cluster overlapping.

So, the classification problem formulation depends on the parameters of classification and these parameters are determined for a problem of classification in a concrete case. The criterion (6) does not depend directly on the parameters. The class of possible solutions of the classification problems depends on the parameters in a concrete case and a unique allotment must be selected from the class on the basis of the criterion (6) for every allotment from the class.

Thus, classification problems can be formulated as follows:

1. Detection of an unknown number c of partially separated fuzzy clusters with given maximal number of elements $1 \leq u < n$ in every class can be considered as the aim of classification.
2. Constructing the allotment $R_c^*(X)$ among unknown number c of fully separate fuzzy clusters with respect to the given minimal value α of tolerance threshold can be considered as the aim of classification.
3. Detection of an unknown number c of partially separated fuzzy clusters such that the number of elements in the intersection area of any pair of fuzzy clusters is less than a priori given value \hat{w} can be considered as the purpose of the classification process.

That is why relational heuristic algorithms of possibilistic clustering which are corresponding to the problem of classification can be developed.

Conclusions

New parameters of relational heuristic algorithms of possibilistic clustering are proposed in the paper. These parameters are theoretical backgrounds for elaboration of new relational heuristic algorithms of possibilistic clustering.

These perspectives for investigations are of great interest both from the theoretical point of view and from the practical one as well.

References

1. Fuzzy Cluster Analysis: Methods for Classification, Data Analysis and Image Recognition / F. Höppner [et al.]. – Chichester : John Wiley & Sons, 1999. – 289 p.
2. Krishnapuram, R. A Possibilistic Approach to Clustering / R. Krishnapuram, J.M. Keller // IEEE Transactions on Fuzzy Systems. – 1993. – Vol. 1 (1). – P. 98–110.

3. Viattchenin, D.A. A Heuristic Approach to Possibilistic Clustering: Algorithms and Applications / D.A. Viattchenin. – Heidelberg : Springer, 2013. – 227 p.
4. Viattchenin, D.A. Direct Heuristic Algorithms of Possibilistic Clustering Based on Transitive Approximation of Fuzzy Tolerance / D.A. Viattchenin, A. Damaratski // Informatica Economica. – 2013. – Vol. 17 (3). – P. 5–15.
5. Dawyndt, P. UPGMA Clustering Revisited: A Weight-Driven Approach to Transitive Approximation / P. Dawyndt, H. De Meyer, B. De Baets // International Journal of Approximate reasoning. – 2006. – Vol. 42 (3). – P. 174–191.
6. Miyamoto, S. Algorithms for Fuzzy Clustering: Methods in C-Means Clustering with Applications / S. Miyamoto, H. Ichihashi, K. Honda. – Heidelberg: Springer, 2008. – 247 p.

USE OF COLOR CHARACTERISTICS AND GRADIENT FIELD DATA IN DISPARITY MAP BUILDING PROCESS

A. Volkovich

United Institute of Informatics Problems of the NAS of Belarus, Minsk

e-mail: anvolkovich@gmail.com

The article discovers the problem of constructing disparity maps for three-dimensional scene reconstruction. The transition to the three-dimensional metric of pixels similarity is considered. The use of the gradient operator to search for and filter uninformative image areas is investigated.

Introduction

The problem of the depth reconstructing on the base of several images captured at the same time, is extremely important and necessary for many applications. One of the areas of practical application is the 3D-reconstructing of the object on its images taken from different angles. This fundamental problem of many researches in computer graphics, but also in cognitive science and neuroscience.

1. Algorithmic model of systems of three dimensional reconstruction

At the present stage of technological development, the industry offers a wide range of standard high-resolution digital and 3D cameras. Manufacturers of optical systems also offer a large variety of lenses with different characteristics and for various purposes that meet the above guidelines.

In developing an artificial system, the processing of binocular information performed by the human brain is implemented as a software-algorithmic system. At the present stage of development of computer vision systems, two approaches to obtain three-dimensional information from several images can be discussed.

The first approach uses characteristic features of depicted objects (object boundaries, corner points, etc.). A disadvantage of this approach is the limited number of points for which the three-dimensional structure is reconstructed. This method gives acceptable results in the case of reconstruction of objects having quite simple forms (architectural scenes, etc.). In the case of more complex objects, this method does not provide the necessary level of detail for highly realistic models; it is used in preparing images for processing by algorithms for the construction of a dense disparity map.

The second approach is based on the use of dense disparity maps and makes it possible to find for each point of the image the corresponding point in the second image or to determine that there are no corresponding pairs.

In general, the three-dimensional model reconstruction from two images can be represented in the following steps:

1. Determination of a number of corresponding points in images. This step can be performed both manually and automatically. Corresponding points allow us to compute the fundamental matrix which describes the geometry of the image pair.

2. There are a number of ways (seven-point algorithm, eight-point algorithm, etc.) to obtain the fundamental matrix, which in turn provides information that helps to limit the search range for conjugate points, reducing it to a search along a straight line, as well as to convert (rectify) the image to enhance search efficiency.

3. Rectification (straightening) of the image based on the fundamental matrix. The straightening of images provides a simpler epipolar geometry of a pair of images, which is subsequently used by many algorithms to construct dense disparity maps for a more efficient search for conjugate points. In rectified images, all epipolar lines all are parallel to the abscissa and the corresponding epipolar lines have the same ordinates.

4. In order to straighten a pair of images, the projective transformation is applied to each them. Since the pair of rectifying transformation is not uniquely determined, projective transformations are selected to minimize the introduced projective distortions. For this purpose, known conjugate points can be used in image rectification methods along with the fundamental matrix.

5. Construction of the disparity map. In computer vision, disparity means displacement of an image point relative to the corresponding reference image point.

6. As a result, a set of conjugate points covering most points of the image is formed, which makes it possible to obtain a fairly detailed model. Some methods of constructing a dense disparity map make it possible to take into account known conjugate points. The reconstruction accuracy largely depends on this step, since the basic set of pairs of conjugate points is formed there.

7. Triangulation. This stage involves determining the coordinates of points in space from information on cameras and known corresponding conjugate points. Note that triangulation can be performed according to a number of methods, including basic ones, that make it possible to create a virtual model (closer--farther) and sophisticated techniques that provide metric reconstruction.

The sequence of steps described above is not compulsory nor universal. The appearance of 3D cameras on the market made it possible to obtain pairs of images perfectly coordinated over the abscissa and thereby eliminated the need for image straightening. The case of creation of a relational model based on preliminarily aligned source data does not require adhering to the actual metric parameters, which also makes it possible to exclude the phase of search for conjugate points.

2. Reconstruction of volume information on stereo images

The main stage of constructing a three-dimensional model is the task of matching point searching on the two rectified images (drawing a map of disparity). In this case the disparity can be understood as the pixel offset of one image point relative to its corresponding point on the other.

Disparity map building is a complex computational problem. The problem of correspondence search completely unresolved. There is a wide range of functions used to determine the similarity measure regions, but as a commonly used input brightness information (monochrome image processing.) The disadvantage of this method is that the multiplicity of interpretations of colors for the image points with the same value of brightness. Thus, in order to improve the adequacy of search results matching seems possible to use color component image pixels. In addition, there is a problem processing homogeneous areas of images that have a small amount of information needed to determine the correspondence between the points of stereo images. Hence the need for a system which allows itself to change the size of a scanning window or screen out non-informative area.

3. The use of color characteristics in corresponding point searching

In world practice, when working with images in stereo reconstruction problems, only intensity information is usually used as a criterion for comparing image points. A disadvantage of this approach is the multiplicity of interpretations of colors for points with the same intensity value. In addition, nonuniformity in the perception of color and monochrome images should be taken into account. This feature is considered in methods of image color model degradation to 256 shades of gray by introducing coefficients applied to the appropriate channels.

Given that most images are originally created in color by a color sensor, in order to enhance the effectiveness of the project, the authors modified the algorithm to enhance its color information processing.

When moving to three components, images can be represented as "clouds" of points in the three-dimensional space with the axes corresponding to color channels of the image. However, the RGB space is not orthogonal because of the specificity of the human's visual analyzer, which has a different number of rods and cones sensitive to a particular color.

Since Euclidean distance, which is applicable to the orthogonal systems, is used as a measure for comparing points in the three-dimensional space, the RGB space should be orthogonalized into the XYZ space. The representation of basic RGB colors as recommended by ITU in the XYZ space is as follows:

$$\begin{aligned} \text{Red: } x &= 0,64y = 0,33; \\ \text{Green: } x &= 0,29y = 0,60; \\ \text{Blue: } x &= 0,15y = 0,06. \end{aligned}$$

After reduction of spaces, operations used in the case of orthogonal systems can be applied to the points.

Thus the measure of similarity of points can be represented as

$$L = \sqrt{(X_2 - X_1)^2 + (Y_2 - Y_1)^2 + (Z_2 - Z_1)^2},$$

where X_n, Y_n, Z_n – are the respective components of compared points.

The maximum length of the vector in the space constructed for the 8-bit color channel is about 441 units, which in terms of informativeness is 1.7 times higher than the comparison of points in a halftone representation of the image with a maximum distance between points of 256 units. In turn, after transition to the color comparison measure, the number of unique values describing points increases by 65 000 times in comparison to monochrome.

Numerical experiments showed a significant improvement in the adequacy of reconstruction by the modified algorithm, which makes it possible to more clearly define the boundaries of objects with similar visual characteristics, objects that visually blend into the background, and objects with fine details.

In turn, there is still the problem of processing of image areas that are not in focus, objects with large homogeneous regions, and an infinitely distant object (for example, the sky). Such image areas have practically no discontinuities (jumps) in intensity that carry the maximum amount of information used in image processing. Obviously, the problem of finding such areas arises. This will make it possible to exclude uninformative areas from processing and, hence, to reduce the number of errors and iterations of the algorithm.

4. Use of a gradient operator to filter out uninformative areas and search-window size determination

Despite an increase in the range of values being compared by using color information remains a problem processing portions of images that are not in focus, objects with large homogeneous regions, as well as infinitely distant objects (such as the sky). Such image area have practically no discontinuities (swings) in the brightness, which in turn carry the maximum amount of information used in image processing. This obviously arises the problem of finding sites that will exclude uninformative treatment area and thus reduce the number of errors and the number of iterations. In order to classify a point located on the intensity jump, the variation in intensity associated with this point should be significantly greater than the intensity variation at the background point. Due to the nature of local computations, the method of determining "significant" values consists in establishing a threshold. In turn, the concept of the first and second derivatives is used to quantify intensity variations.

Determination of the image point as the intensity jump occurs if its two-dimensional derivative of the first order exceeds a predetermined threshold. In accordance with a predetermined connectivity criterion, the connected set of such points is the intensity jump and the prolonged intensity jump is the contour.

One approach to connecting points of the contour is analysis of the characteristics of pixels within a small neighborhood (3×3 or 5×5) of each point (x, y) of the image that has been marked as a contour point (jump point). All points that are similar according to some predefined criteria are connected and form a contour consisting of pixels which satisfy these criteria. In this analysis, the following two main options for establishing the similarity of contour pixels are used, i.e., the gradient operator response and the direction of the gradient vector.

Calculation of the first derivative of the digital image is based on different discrete approximations of a two-dimensional gradient. The direction of the vector gradient coincides with the direction of the maximum rate of change in function f at point (x, y) .

B Calculation of the image gradient consists in obtaining partial derivatives $G_x = df / dx$ and $G_y = df / dy$ for each point. One method of finding the first partial derivatives G_x and G_y at a particular point consists using the following Sobel gradient operator:

$$G_x = (z_7 + 2 * z_8 + z_9) - (z_1 + 2 * z_2 + z_3)$$

$$G_y = (z_3 + 2 * z_6 + z_9) - (z_1 + 2 * z_4 + z_7).$$

For the Sobel operator, which detects horizontal and vertical contours (intensity jumps), it is necessary to determine the appropriate masks for convolution with the original image. It is also possible to change the given formulas so that they provide the maximum response for diagonally directed contours. The sum of coefficients of each mask is zero, so these operators will give a zero response in constant intensity areas, which is characteristic of a differential operator.

Processing of the image by the gradient operator, as well as subsequent binarization of the results, makes it possible to construct the map of calculations.

In most cases occurring in practice, the Sobel operator is applied to a monochrome version of the image. However, as mentioned above, conversion of the color model can distort data in informative areas. Therefore, the authors proposed to process each color channel and their ensemble by the Sobel operator when constructing the map of calculations.

Map computations can be used for image processing using dynamic scanning windows. This process is based on the gradual increase of the window size in the event that it is scanned in the area is insufficient information.

This approach reduces the number of processing errors homogeneous areas, however it should be noted that as the window size is a "swelling" of the boundaries of the object which occurs when a window size 50x50, which distorts the real shape of the object.

Also note averaging the comparison result based on the specific areas of the functions of similarity measures, leading to a reduction in the processing efficiency when the number of pixels in the scanning window begins to exceed the dimension of the range of possible values X Y of the similarity function.

Thus, we can say that it is possible to automatically determine the dimensions of the window in a range that is determined by the bit depth input data based on the requirements of inequality for the converted XYZ color space:

$$S_{scan} \leq \sqrt{\frac{((\max(X) - \min(X))^2 + ((\max(Y) - \min(Y))^2 + ((\max(Z) - \min(Z))^2))}{2}}$$

where S_{scan} – area scanning window;

$\max(X \dots Y \dots Z)$ – maximum value of the component;

$\min(X \dots Y \dots Z)$ – minimum possible value of the component.

In the case where the dimension of the window does not reach the maximum possible "brought together" value in the scanned area sufficient information should make this point screening, due to the fact that information about its position in the space is ambiguous.

Sifting of uninformative areas during computer simulation made it possible to reduce the number of errors in the disparity map and processing time by 20% (including calculation time of local gradient values) compared with the full processing of 860 000 image points.

Conclusion

Investigations in the field of automatic construction of 3D models of real world objects has led to the need to find means of improving the quality of constructing disparity maps. In the course of investigations, the transition to the three-channel metric used to create matching fields in the construction of disparity maps was carried out. Research was also conducted in the field of searching for uninformative areas of images and preliminary sifting of original data leading to obviously erroneous results of calculations.

The set of proposed approaches in conjunction with the actual implementation of the method on a personal computer using GPGPU can contribute to the further development of research in the area of improving the quality of stereo processing in computer vision tasks of real or near-real time.

References

1. Zach, C. Fast Gain-Adaptive KLT Tracking on the GPU / C. Zach, D. Gallup, J.-M. Frahm. – Chapel Hill, NC : Univ. of North Carolina, 2008.
2. Highquality real-time stereo using adaptive cost aggregation and dynamic programming / Liang Wang [et al.] // Proc. Intern. Symp. on 3D Data Processing, Visualization and Transmission. – Chapel Hill, NC : Univ. of North Carolina, 2006.
3. Alyautdinov, M.A. The way to use modern multinuclear processor in neurocomputers to solve mathematical physics problems / M.A. Alyautdinov, G.V. Troepol'skaya // *Neirokomp'yut.: Razrab., Prim.* – 2007. – № 9. – P. 71–80.
4. Volkovich, A.N. Construction of three-dimensional models from images using parallel systems / A.N. Volkovich, D.V. Zhuk, A.V. Tuzikov, // Proc. 9th Intern. Conf. "Pattern Recognition and Information Processing". – Minsk, 2007. – Vol. 2. – P. 232–235.

5. Volkovich, A.N. Construction of three-dimensional models from images using parallel systems / A.N. Volkovich, D.V. Zhuk, A.V. Tuzikov // Proc. 5th Intern. Conf. "Information Processing and Control in Alert Conditions". – Minsk, 2006. – P. 100–104.

6. Systems for automated reconstructing 3D scenes according to several images / V.V. Lyakhovskii [et al.] // Proc. 5th Belorussian Space Congress. – Minsk, 2011. – Vol. 2. – P. 129–133.

GENERAL COVERING ALGORITHM IN SOFTWARE SYSTEM FOR INPUT DATA PREPARATION FOR SINGLE-BEAM VLSI LAYOUT GENERATOR

A. Voronov, A. Doudkin

United Institute of Informatics Problems of the NAS of Belarus, Minsk

e-mail: doudkin@newman.bas-net.by, voronov@newman.bas-net.by

A new software application and algorithm for an automated system of input data preparation of integrated circuit layout generator are proposed. A problem of covering polygons with rectangles is considered. The rectangles must lie entirely within the polygon and it is preferable to cover the polygon with as few rectangles as possible. Functions and a structure of the software are described and given some examples of data processing.

Introduction

There is a problem of formation of layout structures on metallized photo masks by manufacture of integrated circuits, photo-electric converters, LC-indicators, and also many other microelectronic devices [1]. These structures are formed with the help of layout generators. The layout generator builds an integrated circuit (IC) layout on a photo mask from type-setting elements. The type-setting element represents a rectangle. Creation any images of layout structures by means of the generators demands preliminary decomposition of the description of these structures on a set of the rectangles. Thus the number of the rectangles that enclose into the set should be minimal or close to minimal and each rectangle in this set would satisfy to limitations on the sizes. An input information for the generator is formed as the set of rectangles. There are a lot of the automated systems [1, 2, 4–15] of input data preparation for IC layout generators; however they are focused on a limited class of layout objects. For example CATS is a highly scalable and flexible software application that transcribes complex design data into machine readable instructions for e-beam and laser machines used for a pattern generation and manufacturing of IC, MEMS, TFT-LCD, TFH, photonics, and biochip products. CATS has installations in virtually every photomask manufacturing facility worldwide, and is the de facto standard for mask manufacturing, inspection, metrology, and direct-write-on-wafer. But new versions of CATS software do not support an input data format for equipment EM-5109 and EM-5009A2.

A new algorithm and a software system for an automated system of input data preparation for IC layout generator are proposed in this paper. The system realizes multiple document interfaces and uses step-by-step output mode on the screen elements of the found covering. Also there are some specific capabilities in the system that makes a development and an implementation of algorithms for covering more easily. The software system enables to work with a layout description in GDS II format.

1. Layout Decomposition

Different tasks are solved during the input data preparation for IC layout generators, but the main of these ones is layout decomposition or covering layout by rectangles (fig.1).

The decomposition (covering) task is solved for each contour of input data, i.e. a performance of each of these contours as a set of the rectangles is found. Thus each rectangle should satisfy to the given restriction on the minimal length of its side. Special and universal algorithms are used for the decision of a problem of the given stage. Special algorithms solve the covering task of circles, rings, bus-bars and triangles. The universal algorithm solves the covering task of any polygon given by one or several contours. The proposed algorithm is heuristic. It does not guarantee that the received subset of rectangles covers the given contour. Therefore after a reception of the set of the rectangles for covering the task of correctness analysis of this set is solved. For this purpose the rectangles are united in one simply connected contour or a multiply connected contour. If this generated union of the rectangles is the simply connected contour the received covering is correct turns out. If the resulting contour is multiply connected the covering is not correct. Non correct covering is supplemented with the rectangles so that it becomes correct. The obtained sets of rectangles are added in the covering. The prepared input data for IC layout generator you can see on fig.2. Covering task realized not only in our system but also in CATS FLAT Fracturing [4], but the proposed algorithm uses a stage of the covering analysis of correctness due to a special algorithm, that guarantees the correctness of covering.

Further the optimum sequence of the rectangles is founded on the obtained covering set. This sequence is coded using a protocol of an input language for the appropriate IC layout generator. Such task is solved also by software application PG Output Format [4], but a new version of this application does not support GI EM-5109 and GI EM-5009 A2, and our application optimizes input data for such generators.

Thus a number of the rectangles that are enclosed into the result set of rectangles should be minimal or close to minimal, and each rectangle in this set should satisfy to limitations on its sizes. A number of intersections between the rectangles might also be minimized as well as a time for finding covering [8].

The main goal of the software system is to obtain information about input layout patterns in an image format of the generator. The option «Conv» is used to convert graphic images from internal text format into formats MUL, BIT, or PAT, which are an input for layout laser generators EM-5109 and EM-5009A2. To increase the speed up of photomasks production for the generator, all the elements are sorted on a rotation, because a rotary mechanism in generator type EM-5009A2 is slow.

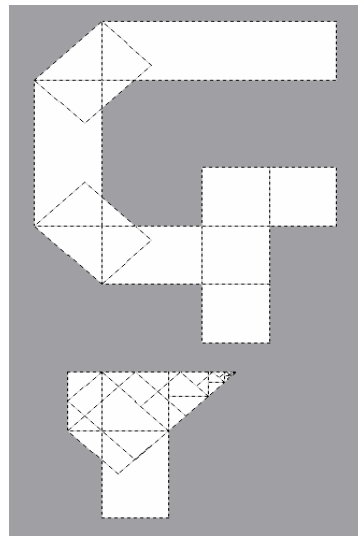


Fig. 1. The result of visualization of output data after covering

One by one we put a description of each rectangle and cover the rectangles by the elements corresponding to the size of an image generator diaphragm (the sizes are from 1 to 300 microns). It is also necessary to determine the type of generator for which information is being prepared (EM-5009 A2 or EM-5109), since the accuracy of the final covering for layout patterns is dependent from the corresponding parameters of image generators. Thus, an error in representation of acute angles and other different elements of the topology is determined by the size of the minimum window of an aperture for image generator. A layout precision must be less than 1 discrete coordinate table (0.125 microns for generator EM-5109 and 0.25 microns for EM-5009A2).

2. General Algorithm and Implementation

The task described above is realized as a set of the appropriate subroutines which are integrated in specially developed visualization system «Polygon». This system essentially facilitates a designer work. It allows executing complex (transparent) layout designing and the analysis of intermediate results. System Polygon realizes the multiple document interfaces and consists of a main window, a menu, a set of floating tool bars and affiliated windows opened as necessary.

Different kind of algorithms is realized in the system, here we describe covering algorithm with analysis stage. Some definitions for the algorithm description are given below.

A *line segment* is a part of a line, if it is bounded by two end points and contains every point on the line between its end points. Examples of the line segments include the sides of a triangle or a square. In geometry a polygon is a flat shape consisting of straight lines that are joined to form a closed chain or a circuit. A *polygon* is a plane figure that is bounded by a closed not crossed itself path, composed of a finite sequence of straight line segments (i.e., by a closed polygonal chain). These segments are called as their edges or sides, and the points where two the edges meet are the polygon's vertices or corners. The simple polygon is single connected if the boundary of the polygon is single and does not cross itself.

Multiply connected polygon W is a plane figure that is bounded by a closed polygonal chains L_1, L_2, \dots, L_g . L_1 is called a main or ground closed polygonal chain and L_2, \dots, L_g are also called contour-cuts. Multiply connected polygon defines the point of the plane located on the borders of representing its contours, as well as the point of the plane inside the main closed polygonal chain, but not inside the contour-cuts. A rectangle belongs to a polygon, if any point of the plane located within or on the border of this rectangle is inside or on the boundary of a multiply connected polygon. The rectangle is h -allowable if the length of any of the parties is not less than some value h , where h is non-zero positive real number. The point of the plane r located inside or on the boundary of the polygon is called the h -covered if there is h -allowable rectangle that belongs to this polygon and the point r lies on the border or inside the rectangle. In the polygon containing acute inside corners, there may be a point located near the acute internal angles, which are not covered by h -allowable rectangles. Moreover, the point of the plane located at the apex of an acute angle is not covered by h -allowable rectangles for any quantity.

A covering of a polygon means a set of multiply h -allowable rectangles satisfying the following conditions: 1) every rectangle of a given covering set belongs to multiply connected polygon, 2) for every h -coverable point r from multiply connected polygon there is at least one rectangle such that the point r lies on the border or inside the rectangle. In this paper we consider the following problem: we must found a correct covering for multi connected polygon W that consists of a minimum or close to the minimum number of h -allowable rectangles.

Example 1. Let's consider a multiply connected polygon W shown in fig. 3. This polygon is given by the main closed polygonal chain L_1 and two contour-cuts L_2, L_3 . The main closed polygonal chain L_1 is the following sequence of vertices defined by its coordinates (2, 2), (2, 12), (7, 12), (9, 10), (9, 9), (7.5, 7.5), (9, 6), (9, 4), (7, 2). Contour-cuts L_2, L_3 are defined by sequences of vertices respectively: (4, 8), (4, 11), (6, 11), (7, 10), (7, 9), (6, 8); (4, 3), (4, 7), (6, 7), (7, 6), (7, 4), (6, 3).

The covering algorithm consists of two stages: construction a previous covering, an analysis this covering on correctness to find holes in covering. Then a special algorithm covers uncovered holes and finds correct covering or an entrance sequence for laser generators.

First stage. Finding previous covering.

The rectangles are constructed on the base of the sides multiply connected polygon. The sides of polygon get over consistently and for every side of the polygon we construct one or more rectangles for previous covering using the special rules. Result of covering you can see on fig. 3, b.

Analysis of the previous covering on correctness.

The above algorithm (at the first stage) for solving the covering problem is heuristic. It's mean that we can't guarantee a valid solution of the covering task. Feature of the previous covering proposed above is that to each side of a multiply connected polygon adjoins one or more rectangles from covering. Exceptions are small areas in sharp internal corners and new contour-cuts added after previous covering. This areas in sharp internal corners can't contain h -covered points. Therefore uncovered regions of a multiply connected polygon after previous covering (except regions that do not contain h -covered points) are the inner regions of the polygon. These internal uncovered regions are the contour-cuts. Farther, these contours-cuts will be called *added after previous covering*. In the next example previous covering is incorrect.

Example 2. On fig. 2, a shows a multiply connected polygon and on fig. 2, b its previous covering after first stage. Obtained covering is incorrect, since it contains two non-covered holes. Then used my algorithm for the analysis previous covering on correctness and find added after previous covering contours-cuts. After analysis we cover contour-cuts added in a result of previous covering by rectangles [16] on fig. 2, c.

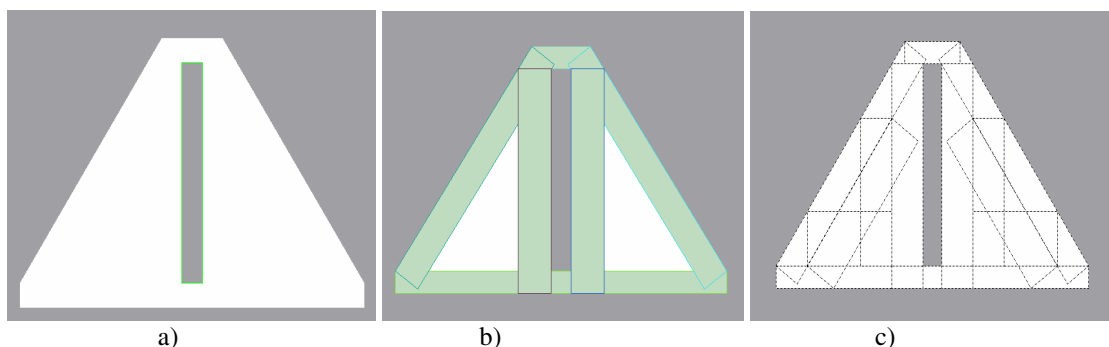


Fig. 2. Covering a) input polygon; b) covering after first stage; c) covering after analysis

This general covering algorithm implemented in system and display good results (table) you can see processing result for this algorithm on generator EM-5109, EM-5009A2, EM-5009M with minimum size 1,2 and 4 micron respectively.

So for 58 % of the polygons previous covering proved correct. As a rule, in orthogonal polygons number of rectangles in covering approximately 2 times less than the number of vertices in the original polygon. This is due to the fact that in the implemented algorithm with previous-covering, added a special optimization procedure. According to this procedure we build next rectangle only on the base of segment that is not a side of previously found rectangles. If the input polygon is similar to the bus chip, the resulting covering contains the same number of rectangles as the vertices in the input polygon. The table 1 shows that the number of uncovered at the first step of searching previous covering contour-cuts are usually constant and does not depend on the minimum size, except only contact bus bar with non-smooth track edges. Increasing the number of rectangles in the covering, usually associated with the complexity of the original polygons: the presence of circular arcs and sharp corners or defects in bus chip. For orthogonal polygons and polygons close to them, which allowed the angles of 45°, the correct covering is formed on the first stage of the covering (the number of rectangles in the resulting coverage is the number of rectangles in the previous covering). For orthogonal polygons, the number of rectangles resulting covering is independent of the minimum size. For arbitrary polygons that contain sharp edges number of the resulting rectangles in covering depends on the applied value of the minimum legal size. With the increase of minimum size decreases the number of rectangles in the resulting covering, which is natural, since the cover sharp corners requires a smaller number of rectangles.

Results of data processing

№	The number of vertices in the original polygon	The number of rectangles in the previous covering			The number of added holes			The resulting number of rectangles in covering			Characteristics of the original polygon
		1 um	2 um	4 um	1 um	2 um	4 um	1 um	2 um	4 um	
1	100	42	42	42	2	2	2	42	42	41	arc
2	65	74	72	72	0	0	0	74	72	72	arc
3	97	104	104	103	0	0	0	104	104	103	arc
4	130	146	144	144	0	0	0	146	144	144	arc
5	50	58	58	56	0	0	0	58	58	56	arc
6	66	74	72	72	0	0	0	74	72	72	arc
7	98	104	104	103	0	0	0	104	104	103	arc
8	50	58	58	56	0	0	0	58	58	56	arc
9	100	104	104	103	0	0	0	104	104	103	arc
10	24	24	24	24	0	0	0	24	24	24	arc

End of the table

№	The number of vertices in the original polygon	The number of rectangles in the previous covering			The number of added holes			The resulting number of rectangles in covering			Characteristics of the original polygon
		1 um	2 um	4 um	1 um	2 um	4 um	1 um	2 um	4 um	
11	10	4	4	4	0	0	0	4	4	4	orthogonal polygon
12	12	7	7	7	0	0	0	7	7	7	bus bar
13	16	37	35	34	0	0	0	37	35	34	bus bar
14	16	10	10	10	0	0	0	10	10	10	arbitrary polygon
15	20	35	34	32	1	1	1	38	37	35	arbitrary polygon
16	25	39	38	36	1	1	1	42	41	39	arbitrary polygon
17	23	87	84	78	2	2	2	91	82	82	arbitrary polygon
18	25	39	38	36	1	1	1	42	41	39	arbitrary polygon
19	18	37	35	33	1	1	1	41	39	37	arbitrary polygon
20	154	188	185	184	0	0	0	188	185	184	arbitrary polygon
21	20	34	32	31	1	1	1	35	33	32	bus bar
22	148	154	152	151	0	0	0	154	152	151	arbitrary polygon

23	155	161	159	158	0	0	0	161	159	158	arbitrary polygon
24	148	186	184	181	0	0	0	186	184	181	arbitrary polygon
25	38	39	39	38	0	0	0	39	39	38	arbitrary polygon
26	176	188	186	184	14	14	14	202	200	198	arc
27	9	46	42	39	6	6	6	52	48	45	arbitrary polygon
28	166	161	161	161	5	5	5	169	169	169	arc
29	16	37	35	34	0	0	0	37	35	34	bus bar
30	16	33	32	32	0	0	0	33	32	32	bus bar
31	16	37	35	34	0	0	0	37	35	34	bus bar
32	23	88	83	77	2	2	2	90	87	81	arbitrary polygon
33	14	34	33	30	1	1	1	38	37	34	arbitrary polygon
34	47	62	60	59	2	2	2	67	65	64	arbitrary polygon
35	9	15	15	13	0	0	0	15	15	13	arbitrary polygon
36	9	16	14	13	0	0	0	16	14	13	arbitrary polygon
37	9	42	40	37	0	0	0	42	40	37	arbitrary polygon
38	168	242	234	225	34	34	34	287	279	270	arbitrary polygon
39	129	216	209	200	31	31	31	256	249	240	arbitrary polygon
40	139	157	156	155	4	4	4	161	160	159	arbitrary polygon
41	164	130	130	130	2	2	2	132	132	132	arbitrary polygon
42	338	286	286	286	3	3	3	295	295	295	bus bar
43	128	121	121	120	6	6	9	122	122	121	bus bar
44	240	227	225	223	27	27	28	296	294	292	bus bar
45	5	13	10	6	0	0	2	13	10	6	arbitrary polygon
46	259	173	167	159	8	8	8	187	181	173	arbitrary polygon
47	183	108	106	105	0	0	0	108	106	105	arbitrary polygon
48	63	49	47	46	2	2	2	55	53	52	bus bar
49	182	107	105	104	0	0	0	107	105	104	arbitrary polygon
50	93	84	80	78	0	0	0	84	80	78	arbitrary polygon
51	127	87	85	83	1	1	1	88	86	84	arbitrary polygon

Table represents results of data processing two thousand two hundred polygons from real projects. To reduce the table we show the every forty-four result, but the table does not significantly reduce impact on the accuracy and reliability.

Conclusion

The proposed software enables to supervise all stages of formation of input sequence for integrated circuit layout generators and process data with usage of technological and technical restrictions of the laser generator. There are alternative variants of decision decomposition task and as you can see in table 1 some algorithms implemented and display good result. All this enables to form better entrance sequences for integrated circuit layout generators.

References

1. Analysis of microlithography in an open architecture TCAD System / V. Axelrad [et al.] // Optical Laser Microlithography IX, Proc. SPIE. – Santa Clara, CA, USA, 1996. – Vol. 2726. – P. 648–659.
2. Technical description and instructions for operating GI EM-5109. – Minsk : Planar, 2001. – 323 p.
3. GDS II format [Electronic resource]. – Mode of access : <http://www.xs4all.nl/~kholwerd/interface/bnf/gdsformat.html>. – Date of access : 16.04.2014.
4. CATS [Electronic resource]. – Mode of access : <http://alt-s.ru/catalog/synopsys/dfm/cats.html>. – Date of access : 16.04.2014.
5. Levenson, M.D. Wavefront engineering from 500 nm to 100 nm CD / M.D. Levenson // Optical Microlithography X: Proc. SPIE. – Santa Clara, CA, USA, 1997. – Vol. 3051. – P. 2–13.

6. Iwai, H. Future semiconductor manufacturing-challenges and opportunities / H. Iwai // IEDM Teach. Dig. – San Francisco, CA, USA, 2004. – P. 1–16.
7. Feinberg, V.Z. Geometrical problems of machine schedulers for great IC / V.Z. Feinberg. – M. : Padio i svyaz, 1987. – 178 p. [in Russian]
8. Current status and future prospect of immersion lithography / S. Owa [et al.] // Proc. SPIE. – San Jose, CA, USA, 2006. – Vol. 6154. – P. 54–65.
9. Matjuskov, V.E. Laser process equipment for precision treatment of materials in electronics / V.E. Matjuskov, S.M. Avakaw // Proceedings of Laser Technology: Indo-Belarus Workshop. – New Delhi, India, 2003. – P. 45–50.
10. Voronov, A.A. A method for covering multiconnected polygons with a set of rectangles / A.A. Voronov, E.A. Shestakov // Informatics. – 2009. – № 4 (24). – P. 39–48.

BASES OF INTELLIGENT SYSTEM CREATION OF DECISION-MAKING SUPPORT ON ROAD-CLIMATIC ZONING

A. Yankovskaya^{1,2,3}, A. Yamshanov³

¹Tomsk State University of Architecture and Building, Russia;

²National Research Tomsk State University, Russia;

³Tomsk State University of Control Systems and Radioelectronics, Russia

e-mail: ayyankov@gmail.com

The paper briefly describes the current state in road-climatic zoning researches and the bases of intelligent system creation for decision-making support and its substantiation. The extended matrix model for data and knowledge representation, the modified algorithms for regularity revealing and pattern recognition are presented. The basis of decision-making and its substantiation, the intelligent system creation and visualization tools including cognitive graphic tools are given.

Introduction

Nowadays, a large amount of intelligent systems (IS) of decision-making support and their substantiation were developed [1–4]. They are based on the different methods and algorithms including the methods of pattern recognition [1, 3, 4]. These systems are used successfully in such problem areas as geology, geo-ecology, building industry, constructional design, medicine, etc.

A large of experience was gained in the field of the IS development for the different problems and interdisciplinary areas in the intelligent system laboratory of Tomsk State University of Architecture and Building by the supervision of A.E. Yankovskaya. Developing such systems in the different problem areas was proposed on the base of the intelligent instrumental software (IIS) IMSLOG [4]. The mathematical apparatus is based on the unusual matrix model of data and knowledge representation (DKR), the test methods of pattern recognition (PR), decision-making and its substantiation (DMIT), cognitive graphic tools (CGT) [5].

However, we do not know about development of the IS for the road-climatic zoning (RCZ) of geographic areas. The fact of such systems creation is beyond question. The solution of the RCZ problem is associated with the necessity to process many environmental and climatic parameters that is typical for large geographical areas in the design and building transport infrastructure objects (TIO). So, this lack of the RCZ problem solution is the main reason of the low reliability level of TIO because of the characteristics underestimation of environment-climatic conditions by designing different territories. The common methodological approach, based on the fundamental research in the field of information and communication technologies, and the IS helps effectively to solve the RCZ problems and to provide the standards improvement of building and design of TIO in the given environment-climatic conditions.

An experience of many of the IS creations, based on the test methods of pattern recognition and the analysis of the RCZ problem, allows to reduce the zoning problem to the pattern recognition one.

Below, the problem area description, the main definition and concepts, the extended matrix model of DKR for the RCZ problem solution are presented. The mathematical foundations of the IS of decision support for the RCZ problem, the visualization and cognitive graphic tools are described.

1. Problem Area Description

It is clear that the knowledge representation for the RCZ problem solution is necessary for qualitative improvement in designing and building of TIO. First, we give the short description to reveal the structure of problem area.

There are four zones on the territory of Russia. Every zone contains such 3 subzones: plain, rolling plain and mountain subzone. Every subzone contains some number of road regions. A number of road regions in a subzone depends on the administrative formation. Zoning is performed on the base of the different characteristic and processes, namely, moisture accumulation characteristics in seasonally frozen ground of roadbed, structures and properties of clayey ground, etc.

The most detailed researches in the RCZ are given in [6-8], which are taken into consideration when structuring the problem area.

2. Basic Concepts and Definitions. Data and Knowledge Representation

The improvement of the matrix model of DKR in the context of problem area under investigation is necessary for the IS creation of the RCZ.

The matrix model, given early in [5], represents data and knowledge as:

- 1) the description matrix Q , which describes objects in the characteristics space of features,
- 2) the distinction matrix R in the classification space of features of 3 types:
 - a) the diagnostic type $R1$ with the included mechanisms of classification when every following column partition previous one into equivalence classes,
 - b) the classification type $R2$ with the independent mechanisms of classification,
 - c) the organization and management type $R3$ with the specified sequence of actions.

Unlike matrix way of DKR was proposed in [5], and where a matrix model was used. There are used such two models: 1) the extended matrix model of DKR on the investigated Western Siberia territory; 2) the matrix model of DKR disregarding the specific of territory based on high-level expert knowledge. We give only difference of the extended matrix model from the model presented in [5] because the framework of the paper would not allow to give the full description of the extended matrix model of DKR in road-climate zoning area.

The extended description matrix Q^e contains 4 additional columns. Every addition column corresponds to compulsory feature which is not used for regularities revealing and decision-making. These columns are used only for the specification of position and border of road-climatic zone, subzone and road region. Ones are numerical values mapped with administrative formation and tracking stations which are used for characteristic feature measurement.

The rows of the extended descriptions matrix Q^e are mapped to tracking stations, columns are mapped to the measured characteristic features which describe environment-climatic conditions and different physical and mechanical characteristic of worked roadbed (geographical complex). The element q_{ij}^e of the matrix Q^e determines the value of characteristic feature j for tracking station i except the values in columns of compulsory features.

An expert fills the distinction matrix R^e simultaneously with Q^e , but R^e is filled by the data from tracking stations under investigation. We will use a type distinction matrix only, which is diagnostic type. Classification features are mapped to zones, subzones and road regions.

3. Revealing Regularities. Decision Making

The modifications of the next algorithms are necessary for the IS creation of the RCZ:

- a) the regularities revealing and building the unconditional irredundant diagnostic test (UIDT) and mixed diagnostic tests (MDT),
- b) feature space optimizing,
- c) selecting optimal subset of the IUDT and MDT on the base of the defined early criteria selection [9],
- d) the DMIS with using different visualization tools including CGT [10, 11].

The subsets of features with particular, easy-to-interpret properties that affect the distinguishability of objects from the different patterns that are stably observed for objects from the learning sample and are exhibited in other objects of the same nature and weight coefficients of features that characterize their individual contribution [5] into objects distinguishability and the information weight on the subset of tests are used for the final decision-making.

These subsets can include constant (taking the same value for all the patterns), stable (constant inside a pattern, but non-constant), non-informative (not distinguishing any pair of objects), alternative (in the sense of their inclusion in the diagnostic test (DT)), dependent (in the sense of the inclusion of subsets of distinguishable pairs of objects), unessential (not included in any IUDT), obligatory (included in all the IUDT), and pseudo-obligatory (which are not obligatory, but included in all the IUDT involved in decision-making) features, and all minimal and all (or part, for a large feature space) irredundant distinguishing subsets of features that are essentially the minimal and irredundant DTs, respectively.

The feature space is optimized on the basis of revealed regularities by: 1) removing all alternative features from every group, except for one; 2) deletion of all dependent features from every group.

To solve the task of object under investigation (OUI) belonging to one or another pattern (to a class at the fixed mechanism of classification) the irredundant implications matrix [5] on the basis of matrixes Q and R is constructed. The irredundant implications matrix sets distinguishability of objects from different

patterns (classes at the fixed mechanism of classification). On the irredundant implications matrix with the application of logical-combinatorial algorithms various regularities are revealed, and irredundant column coverings are constructed [5] prescribing as a matter of fact all IUDT.

Decision-making on belonging of OUI to one or another pattern of each of the IUDT within the framework of the logical-combinatorial approach of the pattern recognition is performed out with use of threshold value of the conditional degree of proximity of the OUI to the considered pattern. This threshold value is calculated on the basis of the admissible error of recognition preassigned by the user [5]. The admissible error is given in percentages.

The definition of a membership object x to pattern k is defined by the coefficient of the conditional degree of proximity (a_k) of an object to the pattern k [5].

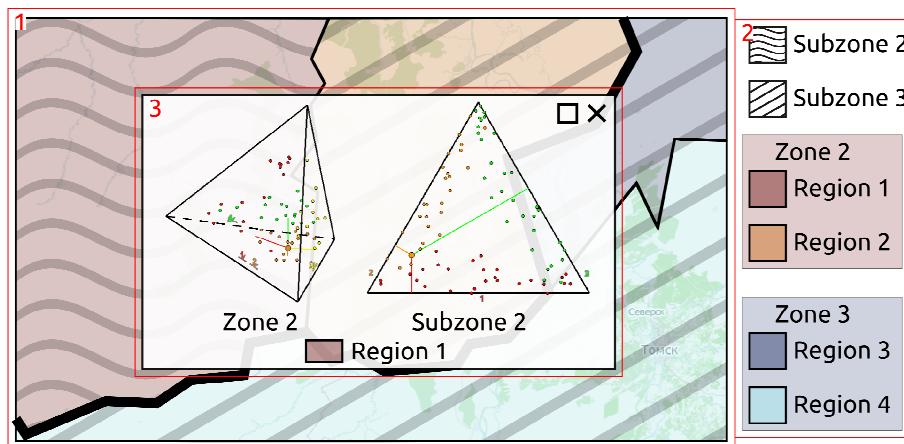
The coefficient of interclass similarity S_k is defined as the ratio of the coefficients sum in pairs similarity (number of the equal values of the same name of features) of objects inside a pattern (class) to number of pairs objects of this pattern (class).

The similarity coefficient S_k^x of an object with pattern (class) is defined as the ratio of the sum of the similarity coefficients of the given object with all objects of the pattern (class) to a number of objects of this pattern (class).

Unfortunately, the framework of paper does not allow even briefly to describe the logical-combinatorial-probabilistic methods of the pattern recognition and parallel algorithms for adaptive recoding features to take into account expert opinions. The modifications of algorithms are related with extending of matrix of DKR.

4. Cognitive Graphic Tools for Decision Making and It's Substantiation in Road-Climatic Zoning Area

One of the most important stages of the IS creation is selection and work of visualization tools to DMIT. There are examples of usage of some visualization tools including CGT [5]. The free-distributed open-street maps (OSM) with information layer overlay for the presentation of common information are proposed. Information layer presents road regions with borders and some information about its. This information is a number of zone and subzone which are determined for road region. The proposed visualization tool is presented in figure.



Visualization tool for representation of the map with zoning results

The information layer is denoted by number 1. It is a transparent layer over the map. The thin black lines separate the different road regions. The different color tones are used for labeling the different zones (red color tone is used for zone 2, blue color tone is used for zone 3). Each color of the road region in every zone is unique color gradation from zone base color given from color transformation in the hue-saturation-bright palette (HSB palette). The wide black lines are used to separate the different zones. Hatching over road region shows subzone type. Only 3 hatching types are used and only 2 types from them presented on figure. Description for all used colors and hatching is presented in the legend (see figure) and it is denoted by number 2. There is a list of all used hatching and presented subzones in the upper part of the legend. The list of all used zones and correlated road regions is in the bottom part of the legend.

The information window for a road region is shown after click on a road region presentation. This window contains full information about a region. This information contains the road region name, 3-simplex and 2-simplex as information about proximity to specific zone (left 3-simplex) and subzone (right 2-simplex). The OUI (road region) is displayed as the circle with a big radius. Objects of learning sample are displayed as circles with smaller radiuses. The distance from the object OUI to an edge is directly-proportional to proximity of the object to the pattern corresponding to the edge. Distances of an OUI to edges are displayed as color lines. Color of an OUI (or objects from a learning sample) is mapped to the pattern which belongs to the concrete object. Mathematical foundations of the visualization of these objects with use of n-simplex are given in [10, 11].

The n-simplex CGT applied to the RCZ allow effectively and clearly to make of decision. The 3-simplex is especially effective at the displaying of patterns of the different objects related to one of the classification features, namely the zones amount of which for Russia today is equal to four.

5. Bases of Intelligent System Creation for Road-Climatic Zoning

Purpose of the IS creating is the fullest account of the environment and climatic conditions which impact on quality of TIO. The system is based on:

- 1) the structuring of data and knowledge in roadbed designing and building field performed cognitologists in together with experts of the RCZ,
- 2) the extended matrix model of DKR,
- 3) the improved parallel algorithms for adaptive features recoding on the base of expert opinions,
- 4) the logical-combinatorial and logical-combinatorial-probabilistic methods of the pattern recognition for regularities revealing and the construction of irredundant unconditional diagnostic tests including the irredundant fault-tolerant mixed diagnostic tests (FT MDT),
- 5) the construction of decision-making rules on each UIDT (MDT or UIDT or FT MDT) and the final decision-making on the base of voting procedure,
- 6) the improvement of cognitive graphic tools for a decision-making support and its substantiation based on the n-simplex and map with information layer overlay.

Previously obtained results of the measurement characteristics of water-thermal regime of roadbed grounds on the territory of the Western Siberia will have been applied for development of data and knowledge base according to dominate geographical complexes [12]. The data and knowledge base will have been developed at the first. This base will contain new and refined data as such as soil evaporation from roadbed, duration of autumn moisture accumulation period, etc.

The matrices Q^e и R^e will be filled as indicated out above. The matrices Q и R will be filled on the base of knowledge of a high-level expert on road building design and building field. The optimal subset of the UIDT will be selected on the base of newly formed subset of criteria. If the fault-tolerant IUDT is necessary, then decision making rules will be constructed for each UIDT and final decision-making will be obtained on the base of the voting procedure on decision making rules set.

Created and improved visualization tool for data structure and revealed regularities representation allow to specialist in road designing and building field to present the data, knowledge base and revealed regularities, and effectively and evidently to make decision and substitute it. The modified algorithms will be implemented on C++ and will be integrated in the IIS IMSLOG [4] as dynamically linked plugins. The IS for RCZ will be constructed on the base of IIS IMSLOG. This system will be used for intelligent technology creation for RCZ.

Conclusion

The extended matrix model of DKR for the IS creation of RCZ is given. This model consists of the extended description matrix Q^e in characteristic features space and distinction matrix R^e in classification features space for each researched administrative formation. The extension is caused by adding unused compulsory features to a set of characteristic features which are not used for regularities revealing. The compulsory features are used the specification of position and the border of road-climatic zone, subzone and road region. The classification features are mapped revealed zone, subzones and road regions for investigated administrative formation. The additional description matrix Q and distinction matrix R are filled on the base of expert knowledge without the specification of territory.

DMIT are performed with the usage of CGT and map with information layer overlay.

The bases of the IS creation of decision making supporting for RCZ are proposed.

Geographical position of boundaries of I-II, II-III and III-IV road-climate zones will have been determined on the base of the developed informational intelligent technology. Therefore, refined map-schema of road-climate zoning including a set of the Russia territories will have been obtained. This map-schema will allow the specialists in the field of road design and building to make reliable decisions. These decisions will allow to ensure the required durability of TIO and to reduce expenditures for TIO construction and its maintenance.

Supported by the Russian Foundation for Basic Research, project no. 14-07-00673, project no. 13-07-00373a and project no. 13-07-98037-r_sibir_a.

References

1. LOREG: Program System of Pattern Recognition and Data Analysis / Y.I. Zhuravlev [et al.] // Pattern Recognition and image Understanding, Infix. – 1999. – P. 244–250.
2. Zagoruiko, N.G. Cognitive Data Analysis / N.G. Zagoruiko. – Novosibirsk : Academic Publishing “Geo”, 2013. – 186 p. [In Russian]
3. Zakrevsky, A.D. Exelor – Expert System of Logic Pattern Recognition / A.D. Zakrevsky // Control Systems and Machines. – 1992. – № 5/6. – P. 118–125. [In Russian]
4. IMSLOG-2002 Software Tool for Supporting Information Technologies of Test Pattern Recognition / A.E. Yankovskaya [et al.] // Pattern Recognition and Image Analysis. – 2003. – Vol. 13 (4). – P. 650–657.
5. Yankovskaya, A.E. Logical tests and means of cognitive graphics / A.E. Yankovskaya. – LAP LAMBERT Academic Publishing, 2011. – 87 p. [In Russian]
6. Sidenko, V.M. About calculated state of roadbed ground base in steppe areas / V.M. Sidenko // Tr. Hadi. – 1957. – Vol 19. – P. 77–82. [In Russian]
7. Vasiliev, A.P. Road designing with accounting of climate impact on road traffic condition / A.P. Vasiliev // M. Transport. – 1986. – P. 248. [In Russian]
8. Efimenko, V.N. Account of region environment-climatic conditions for roadbed design standards / V.N. Efimenko, S.V. Efimenko, M.V. Badina // Science and Technology in Road Area. – 2012. – Vol 1. – P. 15–17. [In Russian]
9. Yankovskaya, A.E. Optimization criteria of irredundant diagnostic tests selection for decision-making in intelligent diagnostic systems / A.E. Yankovskaya // Proceeding of Mathematical Methods of Pattern Recognition. 13th All-Russian Conf. – 2007. – P. 73–76. [In Russian]
10. Yankovskaya, A. Cognitive Graphics Tool Based on 3-Simplex for Decision-Making and Substantiation of Decisions in Intelligent System / A. Yankovskaya, N. Krivdyuk // Proceedings of the IA STED Intern. Conf. Technology for Education and Learning (TEL 2013). – Marina del Rey, USA, 2013. – P. 463–469.
11. Yankovskaya, A.E. Application of Cognitive Graphics Tools in Intelligent Systems / A.E. Yankovskaya, A.V. Yamshanov, N.M. Krivdyuk // IJEIT. – 2014. – Vol. 3 (7). – P. 58–65.
12. Efimenko, S.V. Application of Information System for Road-Climatic Zone Accounting / S.V. Efimenko, D.N. Cherepanov // MGSU Bulletin. – 2013. – Vol. 6. – P. 214–222.

ORAL SPEECH: HOW MUCH INFORMATION IS HIDDEN BEHIND

M. Yerchak

Minsk State Linguistic University, Belarus
yermink@yandex.ru

The idea that any thought is a temporary formation, consisting of images, emotions, and signs is postulated. When expressed in words much of the information fades away: 3-D model is transformed into 1-plane speech expression. Experimental materials, supporting the hypothesis, are presented in the article.

Introduction

Suppose a young man wants to have a cup of coffee. Let's try to imagine how this thought might be represented in his mind. Just like a string of the appropriate words or in some other way? Although the majority of people don't think about the problem at all and many really believe that they think in words still the reality might be quite different. We know that many two-year old children can't speak. But it doesn't mean that they don't think. Most probably they do it using the means they have: sensations, images, emotions and different gestures (signs). These means can effectively be used in the process of thinking later in life as well. The above mentioned young man for instance doesn't need to name the person, wanting coffee because it is obvious for him. And how does he know that he wants coffee? Quite probably thanks to some signals, coming from the appropriate receptors, located in his stomach. As to a cup of coffee, it may be represented in his mind as an image, consisting of the appropriate flavor of black coffee he used to have in the coffee shop near the underground station he passes by on his way home. As we see, the information, preserved in his mind is rather definite, has many aspects and attitudes. When expressed in speech this thought loses much of its richness and specificity.

So the aim of the article is to come to a conclusion what our thoughts consist of, how we can understand them and what factors and conditions stand behind the words, expressing any thought. To be frank we are not always sure that we understand another person's thought correctly and we often say: "What do you mean by saying so"? It's not surprising that some scientists argue that 75% (100%) of the information is rendered in nonverbal form. The complexity of the problem becomes still more evident if we speak about people, belonging to different cultures and speaking different languages.

G. Kaufmann underlines the existence of two paradigms: the symbolist and the conceptualist theories of thinking [1]. The first one includes two main approaches: the imagist and the linguistic. According to the first one thinking is related to reality through imagery: mental images are the primary symbols of thinking; all other symbols are secondary and derived from images. Among the secondary symbols the most important ones are words. Pure verbal thinking does occur too, but it has to be translatable into imagery in order to be understandable. An extreme version of the linguistic theory asserts, that verbal thinking is the only "real" thinking. The conceptualist position affirms that there exist two kinds of processes: a) a mental process where the thought is hatched out; the thought process may accompany speech and imagery but may also go on by itself; b) the second process is mainly needed for conveying a thought; language and imagery are thus placed in a purely external and adventitious position to thinking. So one might conclude that there exists one language that we think in and another language we talk in; no place remains for the possibility that language has a constitutive function in thought. Kaufmann concludes logically that thinking shouldn't be regarded as separate from its symbolic expression, occurring in its own super-physical realm; as for concepts, they should be understood as a seriously misleading maneuver to regard them as some kind of mental objects or entities, as a kind of basic elements of thought in the form of "abstract ideas" or "internal structures". So the problem of thinking (or thought) seems to remain solved only partially and when asked what a thought might consist of we would be in difficulty searching for an answer.

Soviet psychologists define "thought" as the principal element of thinking, aimed at search of connections; it can be expressed in words of a language, but can also resemble associations of meaningful images. Lev Vygotsky underlined the fact that a thought and its expression in speech are different phenomena. A thought doesn't consist of separate words. What is simultaneous in thought is successive in speech. Alexander Sokolov discovered that the electrical activity of the speech organ muscles depended on the difficulty of the proposed task for the subjects; as that type of activation was detected not only in the process of verbal reasoning but also in nonverbal thinking, he concluded that a thought is always linked to a language.

As we can see most scientists agree, that our thoughts might consist of words and images. But there arises the problem with words, because if a thought consisting of words is equal to what is said in oral speech then it's just silent speech. Another problem concerns emotions, that play so marked a role in our life that even the term "Emotional intelligence" was coined to underline its importance. Sometimes emotion is mentioned as something, that tincture image and thought in the flow of consciousness. But there's every reason to believe, that emotion is the third critical component of a thought.

So the preceding material suggests that a thought might be considered as a dynamic temporary formation, constructed of signs (words, phrases), images and emotions in a particular proportion. Some thoughts may, of course, contain only images, others – only emotions, and still others, only words. If they contain only images or only emotions (or both), then the thoughts are practically unconscious and we can't definitely say about what we are thinking. Other thoughts may consist only of signs or mostly of signs, mixed with images and/or emotions. In this case we are already well aware of the language we are using.

Experimental data and discussion

To test the above mentioned hypothesis one needs to have some equipment, that could, in some way, help register the moments, when in a listener's mind there appear images (emotions), stimulated by what is said. Some steps in the direction were already made in the following experiment. Subjects were asked to listen to several short stories (an independent variable) and give signals every moment they had more or less vivid images, stimulated by the contents of each story (dependent variable). Their signals (as well as the story) were being registered on another tape recorder and later on analyzed. It turned out, that in the majority of cases the signals about appearing of images happened to be given when the subjects were listening to the elements of the text, reflecting some change, something new, unexpected etc. Practically the same result was obtained when the subjects were later asked to signal the moments, when they experienced certain emotions. The maximum amounts of signals happened to coincide with those, observed while focusing on images. But as emotions are more difficult (in comparison with images) to realize, the signals, that came from the subjects, were, of course, fewer.

Many other experiments, based on the same methodology, were carried out later and they confirmed the hypothesis. So it was possible to make another step – to see if any thought, expressed by some person, could be analyzed from the point of view, prompted by the hypothesis. So the participants were asked to recollect some thought they expressed to someone or someone's thought expressed to them. When they did it they were asked to specify, whom the idea was addressed, what was there behind the words, that had been present in the thought but didn't get expressed in the utterance.

Let's analyze some of the thoughts that the participants considered worthy of mentioning.

Participant A. "If everything what I had dreamed of once came true, life would be absolutely uninteresting and might have made me commit suicide". The phrase was pronounced during a friendly conversation with a friend; most probably just to reflect a certain aspect of personal experience. Dominance of the sign component (the words here are expressing the person's philosophy or world outlook) seems to be quite evident.

Participant B. "There's no such thing as everlasting love because some day it comes to an end". The judgment was addressed the participant's boy-friend while discussing further development of relations. In spite of the more or less emotional topic of the conversation, this phrase might be regarded as expressing the participant's world outlook and dominance of the sign component seems to be doubtless.

Participant C. "Everything in this world happens in the only possible and the most favorable for us way". The participant expressed the idea to her friend when the latter complained of having abandoned the former place of work as her boss was a willful person. The thought is the person's philosophic generalization and reflection of her optimistic world outlook. Although the participant uses the phrase in order to comfort her friend in some way, still one can see the obvious dominance of the sign component.

Participant D. "Teaching foreign languages on the basis of the computer program "The 25th shot" might be dangerous for a pupil's mind". The thought was spoken out during the lecture on methods of teaching foreign languages after watching some of the programs in action. Here the dominance of the sign component is also evident in spite of possible effect of images the participant had had during the program demonstration.

Participant E. "Time and distance cannot produce any effect on people's feelings if they really and truly love each other. The main thing is the presence of desire to be together". The phrase seems to be not very emotional judging by its contents. But the situation in which it was said shows that the thought is

based on the feeling of anxiety before parting of a dear person. The participant is anxious that long separation might deaden the other person's feelings.

Participant F. "Soon the child is to go to the kindergarten". The dominance of emotional component in the thought becomes evident when the participant informs about her anxiety and fear of the problems that might arise when the child starts visiting the kindergarten.

Participant G. "You know, I have so much to do today and I am so tired". The participant declines an invitation to come to someone's place because doesn't want to see the person. The emotional component is dominant here as well. In fact this is the case when the words don't reflect what is present in the thought of the participant.

Participant H. "It's a wonderful day today". The words are addressed to a boyfriend; the woman participant means to say that it's so nice to be with him that even the day seems to be just wonderful. The emotional component of the thought could have been expressed in many other ways, while the sign component in the situation doesn't matter much.

Participant I. "It's infinite happiness and pleasure to be holding in my arms this little life and feel the beating of the tiny heart". The participant remembered how some days before she had been to church as godmother and had been holding a one-month old baby for 20 minutes and had been experiencing the emotion-thought.

Participant J. "You should have given a hint it was your birthday yesterday". The thought was expressed by the participant's friend who came to her house the following day and saw flowers and presents; he realized it had been the participant's birthday but he had had no idea about it and surely felt ill at ease. One might conclude that image (perception) and sense of guilt served as a basis for the thought in the case.

Participant K. "Why are you smoking in the room? The smoke is spreading all over the apartment and one can't breathe". The phrase was addressed to the participant's brother and meant to blame him with being inattentive to the other members of the family. In this case perception (image) stimulates the emotion and these two components find expression in the above cited words.

Participant L. "The weather is getting worse and worse every day". The participant addressed the words to her mother on their way home from a shop. It was an overcast day and she was cold and melancholy. Evidently the phrase is based on the perception of the day (image) and the appropriate emotion.

All other examples also speak in favor of the hypothesis. Some thoughts really contain only images, others – only emotions, and still others, only words. Some other thoughts are constructed of signs (words, phrases), images and emotions in a particular proportion. The experiments help to see in a different light the role of emotion in thinking. It usually is regarded as a phenomenon, influencing the process of thinking in some way but not as an important component of a thought.

It's interesting to compare the above hypothesis with Johnson-Laird's mental models [2]. He is of the opinion that models can embody abstract predicates that are not visualizable; images, in contrast, represent how something looks from a particular point of view. They are projected from the visualizable aspects of underlying models. A mental model, representing more abstract relations and its associated subconceptual apparatus is likely to be lying behind a scientist's picture of the world.

Conclusion

It was hypothesized that a thought might be regarded as a dynamic temporary formation, constructed of signs (words, phrases), images and emotions in a particular proportion: a 3D-formation. Some thoughts may contain only images, others – only emotions, and still others, only words. Some other thoughts may consist mostly of signs, mixed with images and (or) emotions. To test the hypothesis the participants were asked to listen to several short stories and give signals every moment they had more or less vivid images or emotions. Many more similar experiments were carried out later and they confirmed the hypothesis. So it was possible to go further and ask people to recollect some thought they expressed to someone or someone's thought expressed to them. Then they were asked to specify whom the idea was addressed, what had been present in the thought but disappeared from the phrase. In other words how much information, presented in a thought didn't appear in the appropriate phrase. As we can conclude, a thought may be regarded as a 3D-representation which is to be transformed into a one-plane speech expression.

The proposed materials can be regarded as a cross-cultural study of the language and thought problem. The data can foster mutual understanding between people, belonging to different cultures, as

they are supposed to help discover similarity and difference in thoughts and their expression in speech. And the quantity of the information, that didn't find expression in speech, is worth paying attention to.

References

1. Kaufmann, G. Imagery, Language and Cognition / G. Kaufmann. – Universitetsforlaget, 1980. – 192 p.
2. Johnson-Laird, P.N. Imagery, visualization and thinking / P.N. Johnson-Laird; ed. J. Hochberg // Perception and Cognition at the Century's End. – San Diego, CA : Academic Press, 1998. – P. 441–467.

MULTIVOICE TEXT TO SPEECH SYNTHESIS SYSTEM

V. Zahariev, A. Petrovsky

Belarusian State University of Informatics and Radioelectronics, Minsk

e-mail: {zahariev, palex}@bsuir.by

A scheme of constructing multivoice speech synthesizer based on the use of the synergies of integration of the text to speech and voice conversion systems are presented in this thesis. Such organization of the system allows simultaneous synthesis and modification actions in speech signal, based on an integrated approach to its treatment, significantly reducing the number of errors and artifacts that affect the resulting quality. Applying this approach let to implement a function for multimodal speech synthesizer without significant labor costs for training speech database to add new speakers.

introduction

The key question in the field of the text to speech synthesis (TTS) system researches today is not about providing good levels of the main indicators, for example synthesized speech intelligibility. It is all about more complex features, such as the naturalness of synthesized speech, and support for multiple languages or different speaker voices – multivoice text to speech synthesis systems (MVTTS). The latter aspect requires special treatment and attention, because reconfiguration of the system to the new speaker requires a lot of time and costs of the system developers. The report proposes to consider the possibility of solving the problem of constructing multivoice synthesis system using voice conversion technology, based on the unit selection TTS developed in the speech synthesis and recognition laboratory of the United Institute of Informatics Problems NAS of Belarus [1].

1. Multivoice TTS architecture

Voice conversion is a technique for speech signal processing, which allows to do the transformation of voice parameters characterizing the speech of the source speaker (SS) in the parameters of the target speaker (TS) [2]. Voice conversion as a signal processing technology works with stable over time features of speaker voice, manifested in the speech signal through a change in its acoustic parameters. The attempt to use this technology for TTS to solve the problem of adding multivoice synthesis features is quite obvious. However, the standard solutions in this area, tend to be a very straightforward approach to the application, and can be reduced to a simple piping of two types of systems. In such system, the acoustic processor of the TTS system and the voice conversion system are completely independent: the output signal coming from the TTS is used as input signal for the voice converter. The non-interactive approach has some disadvantages that may result in noticeable quality loss.

The most basic one is that reconstructing the waveform and analyzing it again for converting voices is unessential, taking into account that the same speech model can be used for synthesis and for voice conversion. It is advisable the voice conversion system has access to the speech model parameters and can operate directly on them. Other important limitation is related to the prosodic changes of speech: two different prosodic modifications, from TTS prosody generator and VC system, are performed instead of one, and the consequence is that the quality degradation is higher than strictly necessary. The same phenomenon can occur with durations if the voice conversion system performs any duration modification. Although the unit selection process is optimized for obtaining synthetic speech as natural as possible without significant discontinuities, the fact that the resulting speech signal is to be transformed by means of certain voice conversion function should be taken into account in any way.

Considering the above nuances system architecture allows neutralize the influence of these shortcomings was developed. This is achieved by integrating the conversion module in the acoustic processor unit and rational division of prosody conversion and signal parameters between the two species of systems. Simplified diagram of the architecture is presented in fig. 1 shows an interactive system in which all the limitations commented above are not present anymore.

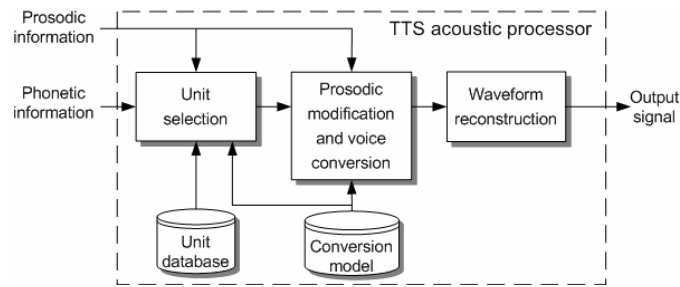


Fig. 1. Architecture of the multivoice TTS

First, voice conversion aspects are taken into account by the unit selector. Second, all the modifications (spectral and prosodic) are performed by a single block so that the signal characteristics are modified only once. Third, the concatenation and reconstruction of the synthetic speech signal are performed after having converted the source voice into the target voice.

2. Speech representation model

In our work we use a model based on STRAIGHT (Speech Transformation and Representation by Adaptive Interpolation of weiGHTed spectrum), originally designed and built by Professor Hideki Kawahara to investigate human speech perception. STRAIGHT uses procedures that can be grouped into three subsystems: a source information extractor, a smoothed time-frequency representation extractor, and a synthesis engine consisting of an excitation source and a time varying filter [3]. The elements that are extracted from the first two subsystems are depicted in fig 2.

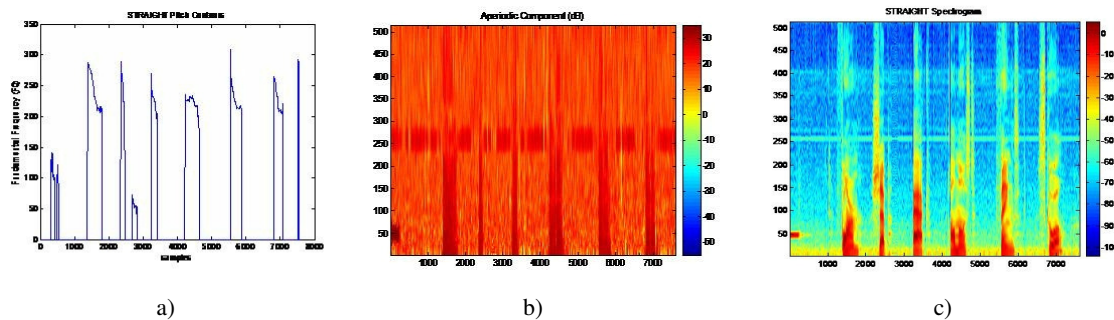


Fig. 2. Parameters extracted by the STRAIGHT model:

a) pitch contour; b) aperiodic component of the speech; c) smoothed spectrogram

Consider the basic nuances of these stages of the model. For the ability to model, adjust, and reproduce speech, it is important to be able to extract F0 trajectories which do not have any trace of interferences caused by the length of the analysis window and the signal waveform. The STRAIGHT system extracts as the fundamental frequency the instantaneous frequency of the fundamental component of the signal. It is extracted as a fixed point of mapping from frequency to instantaneous frequency of a short-term Fourier transform. Normalized autocorrelation based procedure was integrated with the previous instantaneous frequency based procedure to reduce F0 extraction errors further. F0 extraction method assumes that the signal has a nearly harmonic structure. However, there will always exist deviations from periodicity, which introduce additional components on inharmonic frequencies. As such, we can find a measure of aperiodicity by taking the energy on inharmonic frequency normalized by the total energy. The aperiodic component is estimated from residuals between harmonic components and smoothed to generate a time-frequency map of aperiodicity $A(w, t)$ [4].

Second stage, a smoothed time-frequency representation extractor is a heartbeat of the STRAIGHT model. Speech is inherently periodic and, ironically, this does pose problems. Repeated excitation of a resonator is an effective strategy to improve signal to noise ratio for transmitting resonant information in conventional VOCODER systems. However, this repetition introduces periodic interferences both in the time and frequency domains, as shown in the left panel of fig. 3.

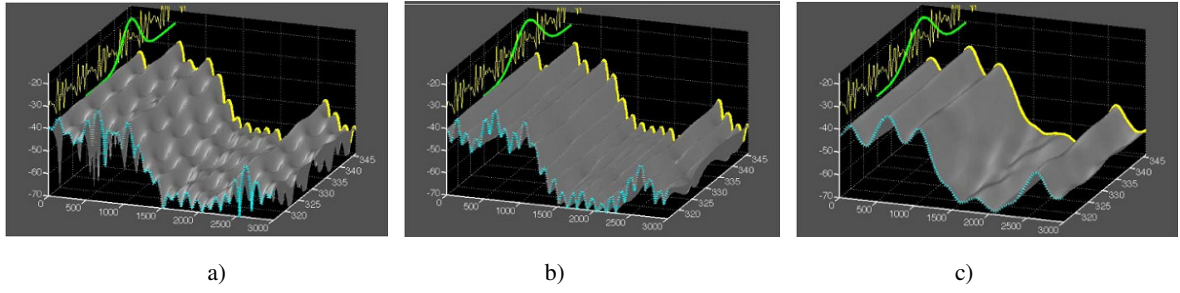


Fig. 3. Smoothed time-frequency representation extraction process:

a) isometric Gaussian window spectrogram; b) reduced temporal variation spectrogram; c) STRAIGHT spectrogram.

It is necessary to reconstruct the underlying smooth time-frequency surface from the representation deteriorated by this interference. The following two step procedure was introduced to solve this problem. The first step is a reduction of the remaining temporal periodicity due to phase interference between adjacent harmonic components is then reduced by introducing a complementary time window. Complementary window $w_c(t)$ of window $w(t)$ is a defined by the following equation:

$$w_c(t) = w(t) \sin \frac{\pi t}{T_0},$$

where T_0 is the fundamental period of the signal. Complementary spectrogram $P_c(w, t)$, calculated using this complementary window, has peaks where spectrogram $P(w, t)$, calculated using the original one, yields dips. A spectrogram with reduced temporal variation $P_R(w, t)$ is then calculated by blending these spectrograms using a numerically optimized mixing coefficient ζ . Fig. 3, b shows the spectrogram with reduced temporal variation $P_R(w, t)$ using an optimized mixing coefficient. Note that all negative spikes found in the top panel, that is $P(w, t)$ disappeared.

The second step is inverse filtering in a spline space to remove frequency domain periodicity while preserving the original spectral levels at harmonic frequencies. STRAIGHT does a smoothing operation using the basis function of the 2nd order B-spline. Because a spectrogram calculated using a complementary set of windows does not consist of line spectra, smoothing kernel h_Q is used to recover smeared values at harmonic frequencies. It is introduced because this operation yields the same results for line spectra and is less sensitive to F0 estimation errors. The following equation yields the reconstructed spectrogram $P_{ST}(w, t)$ (STRAIGHT spectrogram):

$$P_{ST}(w, t) = \left(\int h_Q(\lambda / w_0) P_R^\gamma(w - \lambda, t) d\lambda \right)^{1/\gamma},$$

where w_0 represents F_0 . Parameter γ represents non-linearity and was set to 0,3 based on subjective listening tests. The shape of h_Q is calculated by solving a set of linear equations derived from $w(t)$, $w_c(t)$, ζ and γ . The right panel of fig. 3 shows the STRAIGHT spectrogram of vowel /a/ spoken by a male speaker. Note that interferences due to periodicity are systematically removed from the left to the panel while preserving details at harmonic frequencies [5].

After all kinds of transformations for waveform reconstruction procedure a set of parameters (STRAIGHT spectrogram $P_{ST}(w, t)$, aperiodicity map $A(w, t)$, and F_0 with voicing information $f_0(t)$) are used to synthesize speech. All of these parameters are real valued and enable independent manipulation of parameters without introducing inconsistencies between manipulated values.

3. Voice conversion model

To configure multivoice speech synthesizer on the target speaker is necessary to train the conversion model. This stage consists of the preparation steps of gathering a priori information about the SS and TS voices as phonograms or a set of records of individual units forming a phonetic basis. In our case it is allophones, because they are already presented in the unit database. Given set of phonograms are analyzed and parameterized by the model STRAIGHT. Next, correspondence between these sequences using HMM have to be determined, provided that we have transcripts to TS tracks. Clustering of the resulting joint space parameter vectors using statistical methods, particular based on soft classification of the Gaussian mixture model, is done. Characteristics of these classes or clusters act in the future as the conversion function parameters, which is the core of the conversion model and performs the process of transformation of the voice features. Details about the steps of the training phase were described in the

previous work and literature [6, 7]. Here let us to present the implementation features of the conversion function as one of the main part of the system during runtime.

Conversion function, taken as a core in our work is based on a linear regression model of the first order, using as their coefficients GMM parameters obtained during the training phase. This function has proven effective especially when compared with approaches based on hard clustering parameter space [8]. However, a detailed analysis identified a number of shortcomings. The difficulty of choosing the GMM order with grows of which conversion quality degraded quickly (over smooth effect). And very simple first-order regression model has limited prediction ability. Statistical relationships are present between only one pair of vectors in each i -th time. (fig. 4, a). Model considers only the spatial correlation between the vectors of parameters, excluding the fact that the parameters of the speech signal does not change instantaneously, and have some ergodicity.

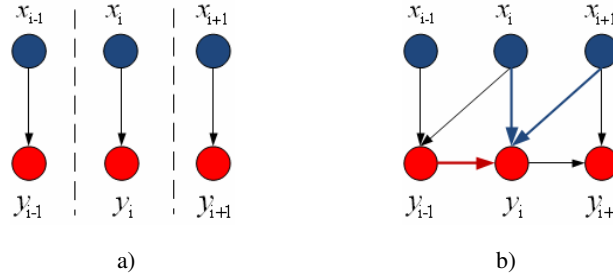


Fig. 4. Types of relationships between pairs of vectors of the training sequence:
a) independent model; b) Markovian process.

In this report we introduce an extended conversion function, based on the GMM and property of ergodicity. That's take into account not only the spatial but also the continual correlation between adjacent vectors for SS and TS, thus giving the sequence properties of Markov process:

$$y_i = \sum_{q=1}^Q p_q(x_i, y_{i-1}, x_{i+1}) [\mu_q + \Phi_q \bar{x}_i^q + \Psi_q \bar{y}_{i-1}^q + \Omega_q \bar{x}_{i+1}^q],$$

where i – frame number $i = 1, \dots, T$, q – number of GMM mixture $q = 1, \dots, Q$, $x = [x_1, x_2, \dots, x_T]^T$, $y = [y_1, y_2, \dots, y_T]^T$ – the sequence of parameter vectors of the SS and TS $x_j, y_j \in \mathfrak{R}^{1 \times p}$, p – the dimension of the parameter vector, $p_q(x_i, y_{i-1}, x_{i+1})$ – posteriori probability that the input vectors x_i and y_{i-1} and x_{i+1} belongs to q -th Gaussian component, $\mu_q = [\mu_1, \mu_2, \dots, \mu_Q]^T$ – vector of the expectations for each component of the mixture of TS $\mu_q \in \mathfrak{R}^{1 \times p}$, $\Phi_q = [\Phi_1, \Phi_2, \dots, \Phi_Q]^T$ – the matrix of regression coefficients for all components of the mixture at variable independent variable x_i $\Psi_q \in \mathfrak{R}^{p \times p}$, $\Psi_q = [\Psi_1, \Psi_2, \dots, \Psi_Q]^T$ – the matrix of regression coefficients for all components of the mixture of first predictor variable y_{i-1} $\Psi_j \in \mathfrak{R}^{p \times p}$, $\Omega_q = [\Omega_1, \Omega_2, \dots, \Omega_Q]^T$ – the matrix of regression coefficients for all components of second predictor variable x_{i+1} $\Omega_j \in \mathfrak{R}^{p \times p}$. Then the problem of finding the unknown parameters $\{\mu, \Phi, \Psi, \Omega\}$ is formulated as optimization problem, the solution of which could be found by the least squares method.

The second conversion model improvement became usage of spectral weighting method, having a deep connection with the physical nature of the speech signal. It allows, on the basis of the available information to make the representation x calculation of the different functions of the frequency scaling for each frame of the input signal as a linear combination of Q basis functions $\{W_q(f)\}$, using the posterior probabilities p_q as weights to classes. Thus, the correct distribution of energy carried by the signal frequencies, and there is no substantial degradation of the resulting envelope. That's help us to solve over smoothing problem.

Also a method for conversion of prosodic features using a parametric representation of the pitch contour was proposed. It is based on the dissent in the fundamental frequency of the singular points, according to the method of Patterson [9]. Next, using a piecewise linear function of a special form, which has different properties in different areas of the pitch frequency values, performed conversion circuit prosodic units. This manipulation provides the best quality conversion of prosodic features without significantly increasing the computational cost.

Waveform reconstruction unit produces the restoration of the speech signal from a set of parameters based on the STRAIGHT model. At the output of the speech signal we get reconstructed phrases with voice characteristics of the target speaker, thereby solving the task of creating MVTTS with text independent learning.

Conclusion

Approach and principle of MVTSS building with the help of VC technology was presented. The rational choice of system architecture based on the interactional approach was provided. The proposed scheme due to synergies from the integration of two types of systems allows to fully exploit the beneficial properties of both, and solves the problem of generating MVTTS with improved exponents for speaker recognition and usability of the system. An extended regression function conversion to improve the accuracy of conversion parameter vectors and methods of application of the spectral weighting for solving the problem of over smoothing was offered. This should raise the quality characteristics of the recovered signal, while maintaining relatively high levels of awareness for the converted signal.

This research was supported by the Belarusian Republican Foundation for Fundamental Research by providing a grant for collaborative research between UIIP NAS of Belarus and BSUIR under contract № 12-1099B from 01.10.2012 named "Voice conversion for text to speech synthesizer".

References

1. Лобанов, Б.М. Компьютерный синтез и клонирование речи / Б.М. Лобанов, Л.И. Цирульник. – Минск : Белорусская наука, 2008. – 344 с.
2. Abe, M. Voice conversion through vector quantization / M. Abe, S. Nakamura, K. Shikano // Proc. of International Conference on Acoustics, Speech and Signal Processing. – N. Y., 1988. – P. 655–658.
3. Kawahara, H. Technical foundations of tandem-straight, a speech analysis, modification and synthesis framework / H. Kawahara, M. Morise // SADHANA – Academy Proceedings in Engineering Sciences. – Vol. 36, № 5. – 2011. – P. 713–722.
4. Kawahara, H. Fixed Point Analysis of Frequency to Instantaneous Frequency Mapping for Accurate Estimation of F0 and Periodicity / H. Kawahara, H. Katayose, A. Cheveigne // Proc. Eurospeech'99. – 1999. – Vol. 6. – P. 2781–2784.
5. Kawahara, H. Restructuring speech representations using a pitch-adaptive time-frequency smoothing and an instantaneous-frequency-based F0 extraction / H. Kawahara, I. Masuda, A. Cheveigne // Speech Communication. – 2004. – Vol. 27. – P.187–207.
6. Moulines, E. Voice conversion: State of the art and perspectives / E. Moulines, Y. Sagisaka // Speech Communication. – 1995. – Vol. 16. – P. 125–224.
7. Stylianou, Y. Statistical methods for voice quality transformation / Y. Stylianou, O. Cappe, E. Moulines // Proc. of European Conference on Speech Communication and Technology. – Madrid, 1995. – P. 447–450.
8. Stylianou, Y. Continuous probabilistic transform for voice conversion / Y. Stylianou, O. Cappe, E. Moulines // Proc. of International Conference on Acoustics, Speech and Signal Processing. – 1998. – Vol. 6.
9. Patterson, D. A linguistic approach to pitch range modeling. PhD dissertation / D. Patterson. – Scotland : University of Edinburgh, 2000. – 201 p.

ALGORITHM OF AUTOMATIC RECONSTRUCTION OF THREE-DIMENSIONAL OBJECTS FOR CAD-SYSTEMS

A. Zakharov, A. Tuzhilkin, A. Zhiznyakov
Murom Institut, Vladimir State University named after
A.G. and N.G. Stoletovs, Russia
e-mail: aa-zaharov@ya.ru

Tasks of automatic reconstruction of three-dimensional objects from orthogonal projections are actual for CAD. This paper summarizes the existing approaches to reconstruction from projections drawing. Algorithm of reconstruction of three-dimensional models based on boundary representation is developed. The algorithm uses patterns for the reconstruction of typical structural elements of the model. Patterns of standard elements are given in the form of adjacency matrix. The developed algorithm processes the data format DXF. The development of software modules are considered. The results of the system are presented.

Introduction

For the moment, a large number of the technical drawings, submitted in both paper and electronic form are accumulated in the archives of the enterprises. An additional point is that the development of many objects often begins with the drawing instead of three-dimensional model usage. Two-dimensional drawings are often difficult to understand, they are inconvenient to update and the ones can not serve as a basis for further developments with computer systems usage. The three-dimensional computer model of the object, which can be used for development of supervisor programs, engineering analysis, visualization, etc., is one of the components of electronic model of product. The modern CAD-systems have the following bevy of tools to generate the three-dimensional models “with a clean slate”: Boolean operations, the operations of object-oriented modeling, 2.5D-operations, modification operations of vertices, edges and faces. Practically all CAD-systems allow the generation of drawings with usage of three-dimensional model. However, generation of three-dimensional model per drawing causes the designers difficulties, related to the lack of software. System fabrication that carries out the automatic reconstruction of three-dimensional models per the technical drawing, would allow many times reducing the time of different purpose objects designing. So it is necessary to develop algorithms and software of automatic reconstruction of three-dimensional models of objects per the technical drawing.

1. Survey of algorithms of three-dimensional reconstruction per technical drawing

All algorithms of three-dimensional models regenerating per drawings can be divided into two groups: B-rep [1–3] and CSG-approaches [4, 5]. CSG-oriented (constructive solid geometry) approach uses “top to bottom” regenerating strategy. The approach is based on the fact that each three-dimensional object can be built with usage of specific two-dimensional primitive in a hierarchical manner. Patterns are found in the drawing, which will serve as a base and they will be used for transformation into three-dimensional model. That done, the designed primitives are collected in the resulting three-dimensional model, using the Boolean operations. The disadvantage of CSG-oriented approach is the fact that with its usage it is difficult to recognize the basic primitives on complex drawings. As well as, it is difficult to imagine the surfaces of complicated shape, when using the CSG-geometry.

B-rep-oriented (bounding representation) approach uses “ascending” technology. B-rep-oriented algorithms generally consist of the following steps:

- generation of possible three-dimensional vertices from drawing;
- edges synthesis per received coordinates of vertices;
- design of faces from edges, lying in plane;
- the formation of three-dimensional object from faces.

Boundary representation provides considerable opportunities for modeling of object complex geometry, it is impossible to reach when using of CSG- approach. However, when using B-rep-representation the more storage space for storage and processing of data is required. An additional point is that, the created model is logically less stable, in other words it is possible to build controversial configurations.

The existing algorithms of reconstruction per drawings are also characterized by the following characteristics: the degree of operator involvement in the reconstruction process, the surfaces geometry, the

number of views on the drawing, the possibility of error correction, the sections processing, etc. The analysis of existing algorithms of three-dimensional reconstruction of the objects models was carried out per technical drawing.

2. The algorithm of automatic reconstruction algorithm on the basis of B-rep representation

Presented algorithm is developed on the basis of B-rep-representation that provides the significant opportunities as to the description of the complex shape geometry. The fundamental concept of the algorithm is to find the structural elements of three-dimensional model with usage of pre-specified patterns. The templates are described by means of matrices. On the basis of the patterns, such elements as holes, slots, chamfers, etc. are described.

The developed algorithm processes the data of vector drawing, saved in DXF (Drawing eXchange Format - format of drawings exchange) format. The algorithm consists of the following steps.

Step 1. Drawing read of vector file. The primitives' parameters (values of vertex coordinates, radius, centers of circles, etc.) are read off.

Step 2. Automatic separation of the drawing per the views. The algorithm operates for the drawings, consisting of three views: front view, top view and view from the left. In this case, the location of the drawing primitives in reference to horizontally and vertically moving straight lines shall be checked. For example, the process of the drawing separation into the main view and the view from the left is finished when there are the primitives that are on the left and on the right from the straight line that is described by the equation $x = A$. At that this straight line does not intersect any primitive of the drawing. In this case, for each vertex of the primitive either condition $x_i > A$ or condition $x_i < A$ shall be met. Similarly, the drawing separation into the main view and top view is carried out.

Step 3 . Finding of the coordinates of the vertices of the three-dimensional model on the basis of view (fig. 1). If the main view to indicate as F (Front), left-side view as L (Left), top view as T (Top), then it is possible to determine the three-dimensional coordinates of the vertices from the following conditions: $x_F = x_T$, $y_F = y_L$, $z_F = z_L$. It is understood that the drawings are designed in the CAD-system with usage of standard tools: object snap, snap to grid, etc. Any inaccuracy in the creation of the primitives leads to an ambiguous interpretation of the image.

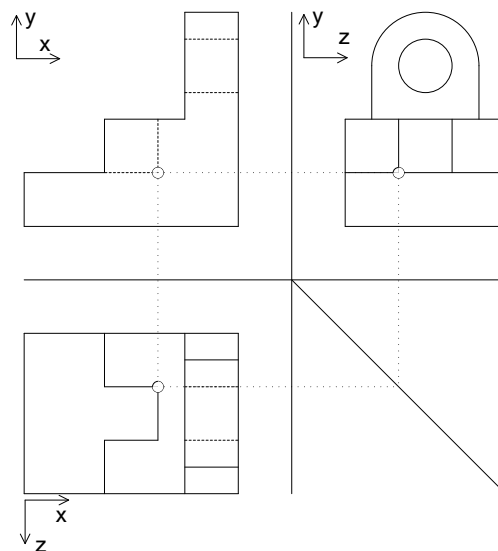


Fig. 1. Conformance of coordinates of the vertices on drawing views

Step 4. Definition and marking of wire model primitives. The wire model consists of the segments of straight lines, arcs, circles, splines, etc. If at least two primitives' projections on views coincide, then there are corresponding to primitive vertices, the coordinates of which are computed at step 3. Description of the primitive in three-dimensional space shall be fed into memory and it shall be assigned with number.

Step 5. Definition and construction of reconstructed object model faces. The face is constructed on the basis of the closed loop of primitives, belonging to the same plane. The closed loop must consist of as low as practicable number of primitives.

Step 6. Definition and construction of structural components of the model on the basis of patterns. Each typical element can be determined on the basis of the adjacency of three-dimensional primitives

(segments of straight lines, circles and arcs), extracted from the drawing. We will assume the primitives that have the common vertex as adjacent ones. Thus, each structural component can be represented by wire model (fig. 2).

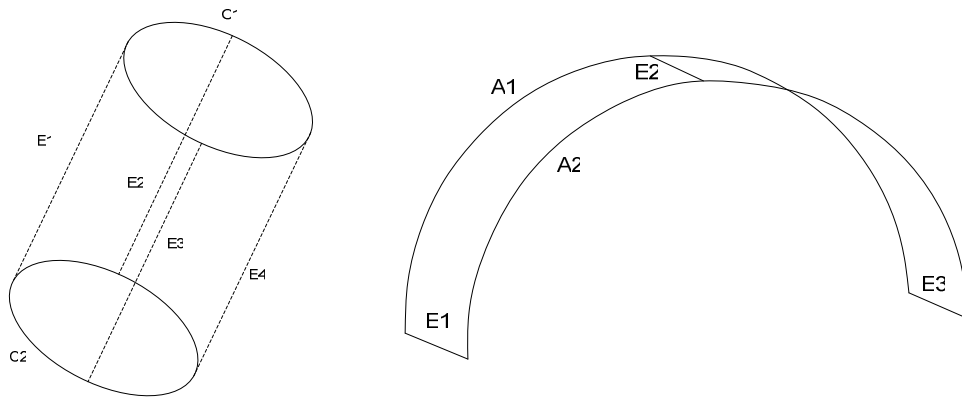


Fig. 2. Structural components, submitted by wire model

The relative positions of the primitives can be described on the basis of the adjacency matrix. Each column describes arc or circle, and the row is the segment of the straight line. The figure that determines both adjacency of primitives and their designation on the drawing is at the intersection of row and column. The matrices that define the patterns of the structural components are saved in the system beforehand. Depending on the relative positions of segment of straight line and circle (or arc) the matrix elements can possess the following values:

- $M[i, j] = 0$, if circle (or arc) and segment of the straight line are not adjacent ones;
- $M[i, j] = 1$, if circle (or arc) and segment of the straight line are adjacent ones. Circle (or arc) and segment of the straight line are represented by unbroken line;
- $M[i, j] = 2$, if circle (or arc) and segment of the straight line are adjacent ones. Circle (or arc) is represented by the unbroken line, and segment of the straight line – by the dashed line;
- $M[i, j] = 3$, if circle (or arc) and segment of the straight line are adjacent ones. Circle (or arc) is represented by the dashed line, and segment of the straight line – by the unbroken line;
- $M[i, j] = 4$, if circle (or arc) and segment of the straight line are adjacent ones. Circle (or arc) and segment of the straight line are represented by the dashed line.

For elements that are shown in the fig. 2, consequently the adjacent matrices will appear as follows:

$$M_1 = \begin{bmatrix} 2 & 2 & K \\ 2 & 2 & K \\ 2 & 2 & K \\ 2 & 2 & K \\ .M & M & O \end{bmatrix}, \quad (1)$$

$$M_2 = \begin{bmatrix} 1 & 1 & K \\ 1 & 1 & K \\ 1 & 1 & K \\ M & M & O \end{bmatrix}. \quad (2)$$

Conclusion

On the basis of the information, extracted from the drawing, the matrices shall be developed and they shall be compared with patterns, stored in the system. If the patterns parameters coincide, then it is possible to conclude as to the presence of one or another structural component. Further, the boundary representation of component form is created.

As test example, the drawing, shown in fig. 1 was taken. The result of reconstruction of views is shown in fig. 3.

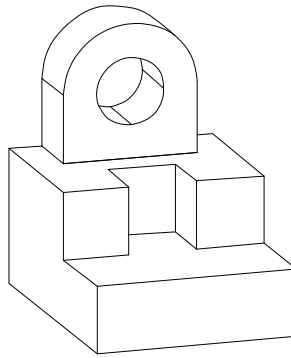


Fig. 3. Reconstructed three-dimensional model

The developed algorithm is fully automatic one and it operates only with the drawings, presented in vector form. The algorithm presents high requirements to the accuracy of drawing construction, the one treats invisible lines and curved surfaces. It is necessary to describe the patterns of all primitives that can occur, when simulating, so that the system would be fully universal for various geometrical shapes. Test modules of the system are developed in the terms of AutoLISP language for AutoCAD system.

This work was supported by RFBR (project № 13-07-97523, 13-07-00825), job number 2014/13 on the implementation of public works in the field of scientific activities of the base part of the state task Russian Ministry of Education

References

1. Dori, D. A Scheme for 3D Object Reconstruction from Dimensioned Orthographic Views / D. Dori, M. Weiss // *Engineering Applications of Artificial Intelligence*. – 1996. – Vol. 9, № 1. – P. 53–64.
2. You, C.F. Reconstruction of curvilinear manifold objects from orthographic views / C.F. You, S.S. Yang // *Computers & Graphics*. – 1996. – Vol. 20, № 2. – P. 275–293.
3. Masuda, H. A cell-based approach for generating solid objects from orthographic projections / H. Masuda, M. Numao // *Computer-Aided Design*. – 1997. – Vol. 29, № 3. – P. 177–187.
4. Golovin, S.I. Automatic reconstruction of curved solids from three orthographic projections / S.I. Golovin, N. Veselov // *Proceedings of SYRCoSE*. – M. : Institute for System Programming of RAS. – 2007. – Vol. 2. – P. 53–58.
5. Wang, Z. Reconstruction of 3D Solid Models Using Fuzzy Logic Recognition / Z. Wang, M. Latif // *Proceedings of the World Congress on Engineering*. – London : Newswood, 2007. – Vol. 1. – P. 37–42.

INTERACTIVE EXTRACTION OF ROADS AND RIVERS IN LOW RESOLUTION OR NOISY SATELLITE IMAGES

B.A. Zalesky, E.N. Seredin

United Institute of Informatics Problems of the NAS of Belarus, Minsk

e-mail: zalesky@newman.bas-net.by

Versions of the interactive polyline active contours and shortest path algorithms to provide practical extraction of roads and rivers in low-resolution and noisy satellite images have been presented. The algorithms are workable tools to facilitate interactive segmentation of the mentioned objects. They are more robust in comparison known versions. Results of experimental study of the algorithms are shown for Landsat 8 and radar satellite images.

Introduction

Extraction of objects in satellite images is an urgent task for many current remote sensing and cartographic applications. We faced this problem when developing software for monitoring of agricultural land on the basis of Landsat 8 multispectral satellite images. Spatial resolution of spectral images turned to be not enough to distinguish roads and rivers from agricultural land. Roads and rivers looked as the surface covered with dense vegetation. Additional tools are needed to find them in panchromatic Landsat 8 pictures having twice the spatial resolution (15 m) compared with the spectral channels (30 m). After, roads and rivers can be excluded from consideration in spectral images. Fragment of a panchromatic Landsat 8 image is shown in fig. 1. Interactive extraction was chosen for that since automatic algorithms do not provide reliable extraction of lengthy objects in a case of low-resolution and noisy satellite images and results of automatic extraction are very difficult to correct manually.

Thereby, the task of interactive extraction of roads and rivers will be understood in the following way. Given a low-resolution panchromatic (for instance, obtained by Landsat 8) or radar satellite image, an expert selects two pixels that restrict fragment of a road or river, which after should be segmented automatically.

Many approaches have been developed to tackle the difficult problem of extraction of roads in satellite images [1, 2], including active contours [3], dynamic programming [4], deformable models [4–6], neural networks [7], simulated annealing and other stochastic processes [8], combinatorial optimization [9, 10] etc. However, almost all of them are either too time consuming or too sensitive to noise and resolution of images in order to be used for interactive segmentation. Besides, the mentioned algorithms require fine parameter tuning or preliminary learning.



Fig.1. Fragment of enhanced panchromatic Landsat 8 image

Really applicable interactive algorithms have to satisfy the following requirements: they should be fast to produce solutions without time delays, be easily and clearly tunable, and be robust to work in presence of distortions and noise.

In accordance with the above requirements two algorithms have been chosen to solve the formulated task. First of them is a new version of active contours algorithm, and the second one is based on finding the graph shortest path.

Both versions allow fast processor and GPU program realizations. The designed algorithms were tested using Landsat 8 and radar images.

1. Extraction of roads and rivers by polylines

The interactive polyline algorithm has been designed to extract road (or river) fragments ending with two pixels, which were selected manually by an expert in a low-resolution or noisy satellite image. Polylines were chosen to provide better control of curvature of the fitting line and prevent deviations

caused by distortions and noise.

Let u and v be end pixels of the current road fragment in an image I . Without loss of generality assume that the pixel u coincides with the coordinate system origin, i.e. $u = (0,0)$. Let k be number line segments in a polyline (fig. 2) with ends u and v . Raster polylines are parameterized by x -coordinates of their pixels, if the angle between the segment \overrightarrow{uv} and the X-axis does not exceed 45° , and parameterizes by y -coordinates otherwise. Denote by $l(x)$ a polyline segment parameterized

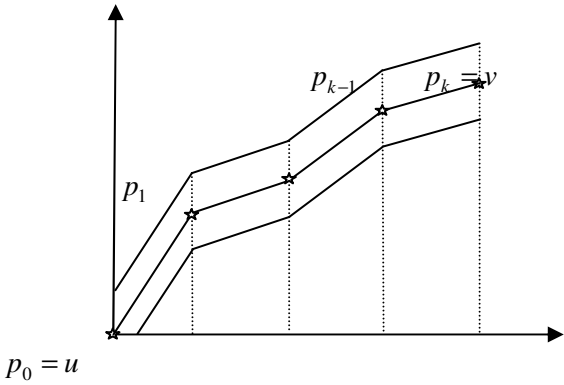


Fig. 2. Polyline model

by x -coordinates of its pixels with ends $u = (u_x, u_y)$, $v = (v_x, v_y)$, knots p_0, p_1, \dots, p_k with x -coordinates $p_{j,x} = j\Delta$, $\Delta = v_x/k$, and angles between adjacent segments $\alpha_1, \dots, \alpha_{k-1}$. For integer $h > 0$ denote by

$$l_h^+(x) = l(x) + h \text{ and } l_h^-(x) = l(x) - h$$

polylines that surround $l(x)$, and by

$$L_h(x) = \{p = (x,y) \mid 0 \leq x \leq v_x, l_h^-(x) \leq y \leq l_h^+(x)\}$$

the set of pixels between lines. Gradient features are used since they allow better extraction of roads from low-resolution images than color ones. The windowed image gradient is denoted by $G=GI$, and its estimate is computed as the convolution $G=I*K$ with user defined windowed gradient kernel K , which depends on road width. The image gradient G is represented by the matrix with vector elements $g(p)=(g_x(p),g_y(p))$, which are indexed by image pixels $p = (p_x, p_y)$. The function F depending on orientation of the gradient with respect to the appropriate polyline segment is taken as a criterion for road extraction. It is of the form

$$F(p_1, \dots, p_{k-1}) = \sum_{\substack{p \in L_h(x): \\ g(p) \neq 0}} \frac{|(p_{j,x} - p_{j-1,x})g_y(p) - (p_{j,y} - p_{j-1,y})g_x(p)|}{\|p_j - p_{j-1}\| \|g(p)\|}, \quad (1)$$

where segment $p_{j-1}p_j$ is chosen so that $p_{j-1,x} \leq p_x < p_{j,x}$, and $\|\cdot\|$ is the Euclidian norm. Actually, summands in (1) are absolute values of cosines of angles between vectors $(-g_y(p), g_x(p))$ that are orthogonal to gradients $g(p)$ and the appropriate segment vectors $\overrightarrow{p_{j-1}p_j}$. Form of the polyline is completely determined by its knots p_1, \dots, p_{k-1} since ends p_0, p_k are fixed by an expert. Therefore, interactive polyline estimate of a road fragment can be built as

$$(p_1^*, \dots, p_{k-1}^*) = \arg \max_{\substack{p_1, \dots, p_{k-1} \\ |\alpha_1| \leq \beta, \dots, |\alpha_{k-1}| \leq \beta}} \{F(p_1, \dots, p_{k-1})\},$$

where the conditional maximum is taken over pixels p_1, \dots, p_{k-1} with coordinates $p_{j,x} = j\Delta$, $0 \leq p_{j,y} \leq v_y$ (or in the general case $r \leq p_{j,y} \leq R$, where r and R are y -coordinates of the rectangle that contains the current road fragment), and predetermined parameter β restricts curvature of the polyline. Evidently, other criteria can be used instead of F . Among them are, for instance

$$F_1 = \sum_{\substack{p \in L_h(x): \\ g(p) \neq 0}} \frac{|(p_{j,x} - p_{j-1,x})g_y(p) - (p_{j,y} - p_{j-1,y})g_x(p)|}{\|p_j - p_{j-1}\| \|g(p)\|^{1-\gamma}},$$

or

$$F_2 = \sum_{\substack{p \in L_h(x): \\ g(p) \neq 0}} \operatorname{acos} \left\{ \frac{\left| (p_{j,x} - p_{j-1,x})g_y(p) - (p_{j,y} - p_{j-1,y})g_x(p) \right|}{\|p_j - p_{j-1}\| \|g(p)\|} \right\},$$

where $0 < \gamma \leq 1$. The function F_1 should be maximized, whereas F_2 – minimized, in order to find the road fragment. Results of experiments that show applicability of the algorithm are places below.

2. Detection of roads and rivers by shortest path algorithms

Combinatorial shortest path algorithms are known tool to solve image processing problems, including roads and river extractions [9, 10], but the task was to adapt the approach for practical interactive extraction of roads from panchromatic images taken with the recently launched Landsat 8 of radar images. Beside that a new way to determine local pixel distances, which are used in shortest path algorithms as arc weights, has been offered and properties of designed algorithms are shown.

Let u and v be end pixels of the current fragment of the road (or river), which were selected by an expert in an image I . Let also the set V consist of pixels of the bounding box B containing the current road fragment. Denote by E the set of arcs that connect adjacent pixels in the bounding box B , where adjacency is considered with respect to the standard 4-pixel or 8-pixel neighborhoods. Then, the pair (V, E) forms a directed graph with starting u and ending v vertices. The shortest path (V, E) problem can be considered after determining weights of the graph.

Experiments have shown application of the 8-pixel neighborhoods provides more adequate results. In the following algorithm description will be done namely for such topology but in the case of 4-pixel neighborhoods principle of the graph construction remains the same.

Remind that $G=GI$ is the image gradient with vector elements $g(p)=(g_x(p),g_y(p))$, which are indexed by image pixels $p = (p_x, p_y)$. Weights of graph arcs should depend on proximity of the current road fragment. The best results were obtained with the following arc weights. For some predefined constant $d > 0$ weights of two horizontal arcs $w_{p,q}$, $q = p \pm (1,0)$ of each vertex p were fixed as

$$w_{p,q} = \begin{cases} d, & \text{if } g(p) + g(q) = 0, \\ d \left(1 - \frac{|g_y(p) + g_y(q)|}{\|g(p) + g(q)\|} \right) & \text{otherwise,} \end{cases}$$

weights of two vertical arcs $w_{p,q}$, $q = p \pm (0,1)$ of each vertex p :

$$w_{p,q} = \begin{cases} d, & \text{if } g(p) + g(q) = 0, \\ d \left(1 - \frac{|g_x(p) + g_x(q)|}{\|g(p) + g(q)\|} \right) & \text{otherwise,} \end{cases}$$

and weights of four diagonal arcs $w_{p,q}$, $q = p + (\pm 1, \pm 1)$ of each vertex p :

$$w_{p,q} = \begin{cases} d, & \text{if } g(p) + g(q) = 0, \\ d \left(1 - \frac{\left| (q_x - p_x)(g_y(p) + g_y(q)) - (q_y - p_y)(g_x(p) + g_x(q)) \right|}{\|g(p) + g(q)\| \sqrt{2}} \right) & \text{otherwise.} \end{cases}$$

After forming the directed graph (V, E) and weight $w_{p,q}$ the shortest path algorithm is used. The trajectory of the shortest path is considered as the estimate of the current road fragment. Results of test of road extraction from Landsat 8 or radar images with help of the offered algorithm turned to be rather robust. One of them is given in the following section.

3. Results of experiments and specificities of algorithms

Large number of experiments has been done for panchromatic and radar images.

The polyline algorithm was tested with number of knots $k = 4,6$ the maximum angle $\beta \leq 30^\circ$, and half the width of the current road fragments $h = 1, \dots, 8$. Results of tests can be seen in fig. 3.

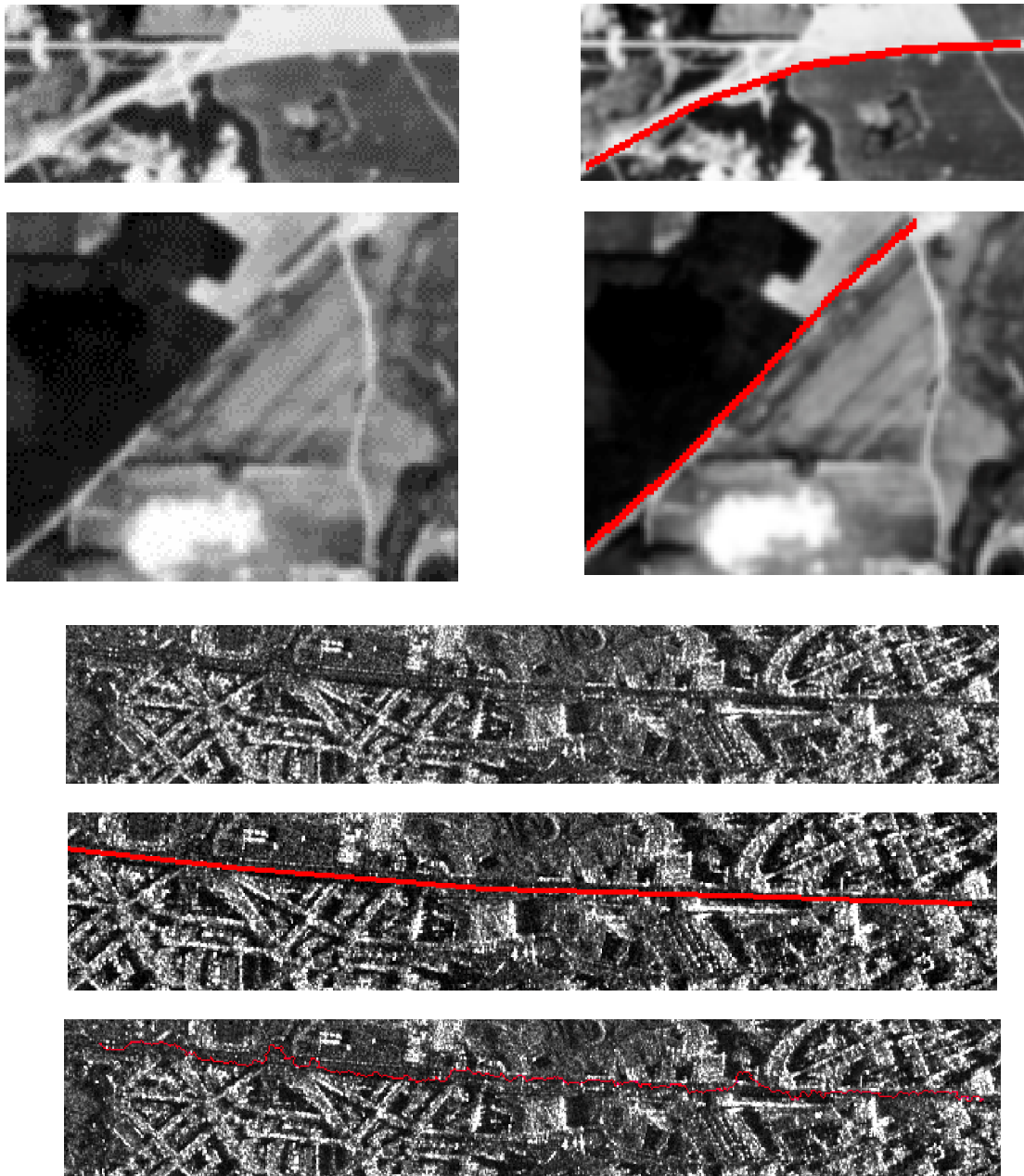


Fig. 3. Road fragments and results of their extraction from panchromatic and radar images by polyline (thick line) and shortest path (thin line) algorithms

Similar results were obtained with help offered versions of shortest path algorithms. Both algorithms have shown practical applicability for extraction of road fragments from low-resolution satellite images.

Conclusion

Versions of interactive polyline active contours and shortest path algorithms have been presented to extract road and river fragments from low-resolution noisy panchromatic and radar satellite images. Algorithms were designed to extract road and river fragments from a satellite images based on fragment ending pixels, which were selected by an expert. Experiments with Landsat 8 images showed applicability of the offered versions that extracted correctly more than 85% of road fragments. In a case of radar images the both algorithms turned to be less accurate because of specificity of radar pictures brightness and other characteristics.

References

1. Supreet, S. Automatic Road Detection of Satellite Images – A Survey / S. Supreet, B. Seema // International Journal of Computer Applications & Information Technology. – 2013. – Vol. 3 (2). – P. 32–34.
2. Kalaivanan, R. Survey on Road Extraction From High Resolution Satellite Images / R. Kalaivanan, S. Mishmala // International Journal of Advanced Research in Computer and Communication Engineering. – 2013. – Vol. 2 (10). – P. 4156–4159.
3. Dailey, M.N. A Family of Quadratic Snakes for Road Extraction / M.N. Dailey, S.M. Dailey, S. Makhanov, K. Honda // Lecture Notes in Computer Science. – 2007. – Vol. 4843. – P. 85–94
4. Dal Poz, A.P. Dynamic Programming Approach For Semi-Automated Road Extraction from Medium- and High-Resolution Images / A.P. Dal Poz, G.M. do Vale // ISPRS Archives. – Vol. 34 (3/W8). – P. 87–91.
5. Urban digital map updating from satellite high resolution images using GIS data as a priori knowledge / T. Bailloeuil [et al.] // Remote Sensing and Data Fusion over Urban Areas, 2nd GRSS/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas. – Urban, 2003. – P. 283–287.
6. Niu, X. A semi-automatic framework for highway extraction and vehicle detection based on a geometric deformable model / X. Niu // Photogrammetric Engineering and Remote Sensing. – 2006. – Vol. 61. – P. 170–186.
7. Automatic Road Extraction from Satellite Imagery Using LEGION Networks // J. Yuan [et al.] // Proceedings of International Joint Conference on Neural Networks. – Atlanta, Georgia, USA, 2009. – P. 3471–3476.
8. Lacoste, C. Unsupervised line network extraction in remote sensing using a polyline process // C. Lacoste, X. Descombes, J. Zerubia // Pattern Recognition. – 2010. – Vol. 43(4). – P. 1631–1641.
9. Color image segmentation: advances and prospects / H. Cheng [et al.] // Pattern Recognition. – 2001. – № 34. – P. 2259–2281.
10. Bucha, V.V. Algorithm Of Interactive Image Segmentation Based On Dynamic Programming / V.V. Bucha, S.V. Ablameyko // Informatika. – 2006. – Vol. 9 (1). – P. 5–15.

Author's index

A

Abaturov V.	6
Ablameyko S.	76, 218, 285
Abramovich M.	11
Alexandrov O.	16
Andrushevich A.	70
Apanasovich T.	157
Apanasovich V.	157, 162
Astafiev A.V.	227

B

Batmunkh Amar	21
Bat-Erdene B.	27
Bausova I.	201, 207
Belotserkovsky A.	213
Berzins G.	201, 207
Bobkov A.	30
Boudnik A.	145
Brazevich H.	168
Britenkov A.	35

D

Damaratski A.	290, 295
Digris A.	162
Dimitrakopoulos G.	40
Dokuchaev A.	35
Dorogov A.Yu.	45
Dorogov Yu.	6
Doudkin A.	65, 305
Dvoenko S.	50, 240
Dzenisiuk Dz.	81

E

Evsutin O.	55
------------	----

F

Faizullin R.T.	60
----------------	----

G

Gafurov S.	30
Ganbat Ts.	27
Ganchenko V.	65
Gladun A.	70
Gulvanskiy V.V.	114
Mashtalir S.	181
Mazouka D.	129
Mikhnova O.	181
Miramontes-Jaramillo D.	187

H

Hancharou D.	76
Hetsevich Yu.	81
Hiuntar A.	81

K

Kanonik D.	86
Kaplun D.I.	114
Karkanitsa H.	94
Karnaukhov V.	187
Kavaliou I.S.	249
Kazlouski A.	89
Kharin Yu.	98
Kharinov M.	103
Kharuzhyk S.	270
Khmelkov A.	213
Khorev I.E.	108
Klionskiy D.M.	114
Kliuchenia V.	119
Kober V.	171, 187
Kolcova I.	223
Kovalev A.	125
Kovalev V.	270
Krasnoproshin V.	30, 129, 135
Krasts J.	201, 207
Kumskov M.	257
Kupriyanov A.V.	231
Kurbatski A.	70
Kuzmich A.I.	140

L

Lazebnikov S.	145
Liauchuk V.A.	152
Lisitsa Y.	157, 162
Lobanov B.	81
Losik G.	168
Lukin A.	176
Lysy S.	81

M

Makovetskii A.	171
Malekzaden M.	191
Mamaev N.	176
Mandzik V.	81
Shakinko I.V.	265
Shangytbayeva G.	254
Shelupanov A.	55
Shestov A.	257

Mitskevich M.	11	Shingaryov I.	162
Moghaddam R.A.	191	Shut O.	260
Moldovan M.	201, 207	Shyrai S.	290
Mortazavi S.M.	191	Sidorenko A.V.	265
Murashov D.	196	Skakun V.	162
N		Snezhko E.	152, 270
Nechval K.N.	201, 207	Starovoitov V.	275
Nechval N.A.	201, 207	Strelchonok V.F.	201, 207
Nedzved A.	76, 213, 285	Svirsky D.	16
Nedzvedz O.	218	Sychyou U.	280
Nguyen Q.T.	135	T	
Nikolaenya E.	290	Tolkachev V.F.	108
Noskov M.	223	Trukhan S.	285
O		Tutatchikov V.	223
Obraztsov V.	135	Tuzhilkin A.	325
Ogorodnikov Y.Y.	60	Tuzikov A.B.	249, 270
Okrut T.	81	V	
Orlov A.A.	227	Viattchenin D.A.	290, 295
P		Vissia H.	30, 135
Paringer R.A.	231	Volkovich A.	300
Petrovsky A.	65, 119, 320	Voronov A.	305
Polyakova G.	235	Voznesenskiy A.S.	114
Popov D.P.	227	Y	
Pshenichny D.	240	Yankovskaya A.	311
R		Yamshanov A.	311
Radoman N.	16	Yaroma A.	295
Rogushina J.	70	Yatskou M.	157, 162
Romanchik V.	30	Yerchak M.	316
Rusak V.	81	Yurin D.	176
S		Z	
Sadykhov R.	89	Zahariev V.	320
Safonau I.	270	Zakharov A.	325
Sarychev A.	145	Zalesky B.A.	329
Semenov D.	244	Zelepugin A.S.	108
Semenov M.	244	Zelepugin S.A.	108
Seredin E.N.	329	Zhiznyakov A.	325
Sergeev R.S.	249	Zhurak M.	98

CONTENTS

Abaturov V., Dorogov Yu. The knowledge extraction analytical platform for embedded industrial applications.....	6
Abramovich M., Mitskevich M. Spatio-temporal cluster analysis of disease	11
Alexandrov O., Svirsky D., Radoman N. Modeling of flow distribution in electric network taking into account the uncertain factors	16
Batmunkh Amar State space representation of multi-input multi-output circuits.....	21
Bat-Erdene B., Ganbat Ts. Computer model for distinguishing Asian people by face.....	27
Bobkov A., Gafurov S., Krasnoproshin V., Romanchik V., Vissia H. Information extraction based on semantic patterns	30
Britenkov A., Dokuchaev A. Structure-functional approach to modulating function trend estimation based on time-series data under uncertain conditions	35
Dimitrakopoulos G. Reconfiguring driving styles based on bayesian networking principles	40
Dorogov A.Yu. Fractal learning of fast orthogonal two-dimensional transformations.....	45
Dvoenko S. Meaningless k -means as k -meanless clustering with the bi-partial approach	50
Evsutin O., Shelupanov A. Creating of the discrete orthogonal transformations with the use of cellular automata dynamics	55
Faizullin R.T., Ogorodnikov Y.Y. Recognition of zero bits of 3-sat problem by applying linear algebra's methods.....	60
Ganchenko V., Doudkin A., Petrovsky A. Decision support for precision farming complex	65
Gladun A., Rogushina J., Andrushevich A., Kurbatski A. User-oriented recognition of intelligent information objects in distributed dynamic informational web-space.....	70
Hancharou D., Nedzved A., Ablameyko S. Skeletonization algorithm of high resolution vascular data	76
Hetsevich Yu., Mandzik V., Rusak V., Hiuntar A., Okrut T., Lobanov B., Lysy S., Dzenisiuk Dz. The system of generation of phonetic transcriptions for input electronic texts in Belarusian	81
Kanonik D. Distinguishing mental activity types by electroencephalography records.....	86
Kazlouski A., Sadykhov R. Plain objects detection in image based on a contour tracing algorithm in binary images.....	89
Karkanitsa H. Development of domain-specific language for constructing dynamic subject domains	94
Kharin Yu., Zhurak M. Statistical analysis of spatio-temporal data	

by poisson conditional autoregressive model.....	98
Kharinov M. Hierarchical pixel clustering for image segmentation.....	103
Khorev I.E., Zelepugin S.A., Tolkachev V.F., Zelepugin A.S. Image analysis in high-velocity interaction of a group of bodies with a target	108
Klionskiy D.M., Kaplun D.I., Voznesenskiy A.S., Gulvanskiy V.V. Study of digital filter banks and their software-hardware implementations for wideband monitoring	114
Kliuchenia V., Petrovsky A. Architecture of the DCT-IDCT processor for lossless scheme coding.....	119
Kovalev A. An agent-based model of German-French inter-border residential mobility in Strasbourg region	125
Krasnoproshin V., Mazouka D. High-level rendering pipeline construction tools.....	129
Krasnoproshin V., Obratsov V., Vissia H., Nguyen Q.T. Data mining and knowledge-based technology	135
Kuzmich A.I. Standartization of mobile heterogenous objects monitoring based on pattern recognition	140
Lazebnikov S., Boudnik A., Sarychev A. Application of big data technologies to image processing	145
Liauchuk V.A., Snezhko E.V. A method for coarse segmentation of pathological lungs on CT images using ribcage	152
Lisitsa Y., Yatskou M., Apanasovich V., Apanasovich T. Studying the advanced segmentation methods with the computer-simulated images	157
Lisitsa Y., Yatskou M., Skakun V., Digris A., Shingaryov I., Apanasovich V. Simulation model to study denoising methods.....	162
Losik G., Brazevich H. Brain coding of information on the object shape by its physical model.....	168
Makovetskii A., Kober V. Frequency analysis of gradient descent metod for imagerestoration	171
Mamaev N., Lukin A., Yurin D. HENLM-3D: A 3D computer tomography image denoising algorithm.....	176
Mashtalir S., Mikhnova O. Video content analysis and key frame extraction method.....	181
Miramontes-Jaramillo D., Kober V., Karnaukhov V. Fast image matching with sliding hogs	187
Moghaddam R.A., Mortazavi S.M., Malekzadeh M. Reducing energy consumption in wireless sensor networks by CICA algorithm in comparison with the genetic algorithm	191
Murashov D. Application of structure tensor features for describing an artistic manner.....	196

Nechval N.A., Nechval K.N., Bausova I., Berzins G., Krasts J., Strelchonok V.F., Moldovan M. Distance-based embedding procedures for pattern classification	201
Nechval N.A., Nechval K.N., Bausova I., Berzins G., Krasts J., Strelchonok V.F., Moldovan M. Statistical validation of simulation models	207
Nedzved A., Belotserkovsky A., Khmelkov A. System for automatic parallelisation of processing large volumes of images using multi-user resources	213
Nedzvedz O., Ablameyko S. Detection patterns of living cells from its aggregation on the digital image.....	218
Noskov M., Tutatchikov V., Kolcova I. Three-dimensional Fast Fourier Transform algorithm modification by analogue of Cooley-Tukey	223
Orlov A.A., Astafiev A.V., Popov D.P. Development of algorithm for localization of symbolic with the use of analysis of the color map	227
Paringer R.A., Kupriyanov A.V. Research methods for classification of the crystallograms images.....	231
Polyakova G. Logical-and-probability simulation model of data processing and analysis.....	235
Pshenichny D., Dvoenko S. Elimination of the non-positive definiteness of matrices of pairwise comparisons.....	240
Semenov M., Semenov D. Targets visualization tool to explore testing results	244
Sergeev R.S., Tuzikov A.V., Kavaliou I.S. Mutation analysis of genetic sequences and building mutational database in context of tuberculosis treatment	249
Shangytbayeva G. Planning of the enterprise resources.....	254
Shestov A., Kumskov M. Multiscale molecular surface descriptors.....	257
Shut O. The complexity of inductive resolution algorithms in pattern recognition problems.....	260
Sidorenko A.V., Shakinko I.V. The modified principal component analysis of information encrypted using deterministic chaos	265
Snezhko E., Kharuzhyk S., Tuzikov A., Kovalev V., Safonau I. A scheme for nodules detection on CT images of lungs.....	270
Starovoitov V. Image processing for fundus image analysis	275
Sychyou U. Method of mobile-robot control to implement search motion.....	280
Trukhan S., Nedzved A., Ablameyko S. Region growing segmentation of CT-image by the analysis of 3D local neighborhood	285
Viattchenin D.A., Damaratski A., Nikolaenya E., Shyrai S. An outline for an approach to automatic labeling for interpretation of heuristic possibilistic clustering results	290

Viattchenin D.A., Damaratski A., Yaroma A. New parameters for relational heuristic algorithms of possibilistic clustering	295
Volkovich A. Use of color characteristics and gradient field data in disparity map building process.....	300
Voronov A., Doudkin A. General covering algorithm in software system for input data preparation for single-beam VLSI layout generator	305
Yankovskaya A., Yamshanov A. Bases of intelligent system creation of decision-making support on road-climatic zoning	311
Yerchak M. Oral speech: how much information is hidden behind	316
Zahariev V., Petrovsky A. Multivoice text to speech synthesis system	320
Zakharov A., Tuzhilkin A., Zhiznyakov A. Algorithm of automatic reconstruction of three-dimensional objects for CAD-systems	325
Zalesky B.A., Seregin E.N. Interactive extraction of roads and rivers in low resolution or noisy satellite images	329
Author's index	334

**РАСПОЗНАВАНИЕ ОБРАЗОВ И ОБРАБОТКА ИНФОРМАЦИИ
(PRIP'2014)**

Доклады XII Международной конференции
28–30 мая 2014 г., Минск, Беларусь

На английском языке

Ответственный за выпуск Н.А. Рудя

Подписано в печать 05.05.2004. Формат 60×84 1/8. Бумага офсетная. Гарнитура Таймс.
Усл. печ. л. 39,5. Уч.-изд. л. 32,9. Тираж 100 экз. Заказ № 16.

Издатель и полиграфическое исполнение:
Государственное научное учреждение «Объединенный институт проблем
информатики Национальной академии наук Беларуси».
Свидетельство о государственной регистрации издателя, изготовителя,
распространителя печатных изданий
№ 1/274 от 04.04.2014 г.
Ул. Сурганова, 6, 220012, Минск