Towards Creation of SmileID Obtained from Face Biometrics Binded to Concantenated Error-Correcting Codes

Boris Assanovich
Yanka Kupala State University of Grodno
Grodno, Belarus,
bas@grsu.by

***Abstract.* A new biometric system based on smile imprint obtained from face video frames with the use of stacked autoencoder and fuzzy commitment scheme exploiting the concatenated Reed-Solomon and linear error-correcting codes has been proposed. The performance of system verified with smiles from UvA-NEMO database has shown the achievement of FRR=0.5-1% for data size of 31-63 elements.**

***Keywords*: autoencoder, biometric features, smile imprint, fuzzy commitment, concatenated Reed-Solomon codes**

I. Introduction

A human smile that is a key factor in determining person's psychological state, can also be used as a behavioral biometric identification element. Recently, several applications have appeared that implement the concept of SmileID [1] related to the face biometrics where identity is verified remotely. Last years, facial biometry has become one of the most preferred biometric methods both in video surveillance and in banking [2] due to the fact that it does not require precision equipment and uses a non-contact data processing method.

Few years ago Murat Taskiran et all [3] applied dynamic face features extraction from videos and used them for face recognition. Authors performed the analysis of face videos and extracted the statistical properties of facial distances during several phases of spontaneous and posed smiles on the UvA-NEMO smile database, that have been created before for biometric applications by Dibeklioğlu et all [4].

Despite the fact that there is a number of techniques that use facial dynamics to identify a person with extracted from face various parameters, the deep learning methods are increasingly being exploited for recognition tasks. Classical linear methods of image processing and feature extraction based on principal components analysis (PCA) are replaced by non-linear transformations. Compared to PCA, the use of an autoencoder increases significantly the classification accuracy, especially with a large number of items [5].

In this paper, we will consider the use of an autoencoder to extract soft biometric information from a person's smile and apply them for authentication to provide access to digital services. In addition, we will propose an error correction method for creating biometric templates with the use of Juels and Wattenberg (JW) fuzzy commitment scheme [6] that can be revoked if compromised.

The article demonstrates the creation of a digital face *smile imprint* to provide a secure biometric interface for organizing access to digital services. It continues in Section 2 by reviewing the related structure of an autoencoder and the error correcting code parameters. The methodology used for the research work is presented in Section 3. Section 4 describes the results of this study. The paper concludes in Section 5 and also provides some direction for future work.

II. Autoencoders and Error-Correcting Codes

A. Autoencoders

There are several types of autoencoders. Sparse Autoencoder has a dimension of the hidden layer that is greater than the input. It consists of two parts: coder (encoder) *G* and decoder *F* [7]. The encoder translates the input signal into its representation (code): $y = G(x)$, and the decoder restores the signal by its code: $x' = F(y)$. Moreover, the transformation functions *F* and *G* contain activation function, weights and biases of trained artificial neural networks. By changing the mappings *F*, *G*, autoencoder tends to find out the identity function $x = F(G(x))$, minimizing the kind of error based on some functional $L = (x, F(G(x')))$. Let consider that a vector $x \in \mathbf{R}$ connected to the input of an autoencoder. Then the encoder maps the vector *x* to another vector $y \in \mathbf{R}$ as follows $y = h^i(W^i x + b^i)$ where the superscript *i* indicates the *i*-th layer. Then h^i is a transfer function for the encoder, $W^i \in \mathbf{R}$ is a weight matrix, and $b^i \in \mathrm{R}$

Pattern Recognition and Information Processing (PRIP'2021) : Proceedings of the 15th International Conference, 21–24 Sept. 2021, Minsk, Belarus. – Minsk : UIIP NASB, 2021. – 246 p. – ISBN 978-985-7198-07-8.

is a bias vector. Hence, the decoder maps the encoded representation *y* back into an estimate of the original input vector *x*, as follows: $x' = h^{i+1}(W^{i+1}x + b^{i+1})$ where the superscript *i*+1 represents the $i+1$ layer. Then a transfer function h^{i+1} for the decoder has a factor $W^{i+1} \in \mathbf{R}$ that is a weight matrix, and $b^{i+1} \in \mathbf{R}$ is a bias vector correspondently. If the encoder has only two layers then the expression for the transfer function can be represented as $x' = h^2(W^2 y + b^2)$. In our setup we applied the so-called stacked autoencoder (SAE) that is a neural network including only 2 layers where output of each hidden layer is connected to the input of the successive one. The effectiveness of user comparisons depends on similarity rates, which are often determined by the distribution of root mean square (RMS) distances of their characteristics.

The more the two distributions are separated and the smaller the standard deviation for each distribution, the better is the separation of the classified classes. This property of distributions is estimated by such a parameter as decidability index

$$DI = \frac{|\mu_g - \mu_i|}{\sqrt{\sigma_g^2 + \sigma_i^2 / 2}}, \quad (1)$$

where μ_g, μ_i and σ_g, σ_i are the means and standard deviations of genuine and imposter distributions.

In addition to the decidability index, an equal error rate (ERR), which is the rate at which a false accept rate (FAR) is equal to a false rejection rate (FRR), is normally used as a measure of biometric system (BS) verification accuracy. In biometrics FAR is the rate at which an imposter print is incorrectly accepted as genuine and FRR is the rate at which a genuine print is incorrectly rejected as imposter.

The use of autoencoders will make it possible to present compactly biometric features and then apply the scheme to organize Human Computer Interface (HCI), where instead of tokens and passwords, a biometric key can be exploited for authentication. To handle the variability inherent in biometric verification, it is necessary to create and store a template for each user. To create it, we will use the fuzzy commitment scheme with application of error-correcting codes (ECC).

B. Error-Correcting Codes

In this paper, in contrast to the generally accepted application of binary ECC in Biometric System, we consider the use of non-binary Reed-Solomon codes (RS). However, the parameters we put for evaluating the effectiveness of biometric systems can easily be used for binary ECC.

More formally, a non-binary ECC have codewords *C* of a certain length *n*, consisting of symbols belonging to the set *C**∈{0, *q*-1}, where *q* is some integer. The use of the code is aimed at encoding a message of length *k* with the addition of redundancy *r*=*n*-*k*, so that if a certain number of symbols is corrupted, it is still possible to get the correct codeword *C* and message *R*. The robustness of an ECC depends upon some distance between codewords. Let the RS code be defined over the Galois Field $\mathrm{GF}(2^m)$ with a redundancy of *n*-*k* symbols and let the Symbol Error Rate (SER) caused by fuzziness of biometric data or so called “biometric noise” be *p*. The important performance parameters to consider when using EEC in BS are still FAR and FRR. The FRR, which depends on SER of RS code, is actually can be upper bounded by the probability that more than $(n-k)/2$ errors occur, i.e. [8]:

$$FRR \le \sum_{i=\lfloor (n-k)/2 \rfloor + 1}^{n} \binom{n}{i} p^i (1-p)^{n-i} \approx (np)^{\lfloor (n-k)/2 \rfloor + 1}, \quad (2)$$

where *n*, *k* are the ECC parameters defined above, and *t* is its error correction capacity, i.e. the ability to correct any set of up to *t* symbol errors.

Considering that an imposter produces a random syndrome during decoding for uncorrectable error patterns, and it is accepted by BS, FAR will be the probability that it is valid [8], i.e.

$$FAR = \sum_{i=0}^{\lfloor (n-k)/2 \rfloor} \binom{n}{i} (n+1)^{-(n-k)} \approx q^{-(n-k)/2} \quad (3)$$

From (2), (3), we see that FRR, FAR can be reduced by increasing (*n*-*k*) and *t*.

The dimension of the $\mathrm{GF}(2^m)$ field and the redundancy of the RS code significantly affect the length of the cryptographic key *R* in the biometric JW scheme. The use of non-binary RS codes has the advantage that an increase in the symbol dimension leads to an increase in the length of their bit representation. On the other hand, in order to obtain high error correction capacity *t* of ECC, it is necessary to increase the redundancy, which reduces the code rate *k/n* and leads to the need to use several ECC codewords. For example, to obtain a user's secret key *R* with a length of more than 128 bits, when using RS (31,9) code over $\mathrm{GF}(2^5)$ with a code rate *k/n* of about 0.3, only 45 bits can be placed in one codeword. Hence for the key length of more than 128 bits, three such codewords are required. Whereas when using the RS (63,15) code with the rate 0.24 over $\mathrm{GF}(2^6)$, the key length of 180 bits can be distributed between two codewords of this RS code.

To evaluate the effectiveness of these codes, the corresponding FRR values for these codes have been calculated at different symbol error probabilities and placed in Table I.

TABLE I. EVALUATION OF FRR FOR RS CODES (63,15), (31,9)

SER	FRR: RS (63,15)	FRR: RS (31,9)
0.0050	2.8692e-13	1.9230e-10
0.0100	9.6275e-06	7.8766e-07
0.0150	0.2431	1.0220e-04

It follows from Table 1 that with the increase of *p*, FRR grows exponentially and, taking into account (2), to reach FRR=1.0e-04 for RS code (63,15), it is necessary to reduce SER by 1.36 times. Thus, in order to achieve the required performance of the biometric system, it is relevant to introduce a significant redundancy by ECC applied. In this case one of the ways to increase the efficiency of error-correcting coding is the transition to concatenated ECC. It is a class of error-correcting codes derived by combining an inner code and an outer code that can be tuned in a given way and show better performance than ECC of a certain type. According to our estimates, to reduce SER by several times, it is sufficient to use a class of linear codes with a suitable length.

III. Proposed System

In this paper, we have used an autoencoder to obtain the biometric data on a person’s smile and bind it to a secure user key. Due to the fact that biometric data has instability, error-correctiing codes should be adopted to ensure that fuzziness of biometric data can be alleviated.

In this study we propose to apply the concatenated ECC based on non-binary RS codes and binary linear codes with the use of hard-decision decoding technique. Initially, the user key will be encoded with a non-binary RS code, and then the bit representations of the symbols will be additionally encoded with a linear binary code. Then, after eliminating “biometric noise” in the received code vectors, the key is extracted by decoding the code constructions in reverse order.

The model of the proposed system is depicted in Fig. 1. Wel introduce the term *Smile Imprint* of biometric data obtained from SAE output layer and concatenated $Y = y_1 \cup y_2 ... \cup y_M$ to form a supervector from several vectors y_i, where *M* is a number of processed frames from “smile signal”. It should be noted that the *Preprocessor* block also performs such operations as smile detection and smile frames selection related to its main three phases (onset, apex, offset) as it is also described in [9].

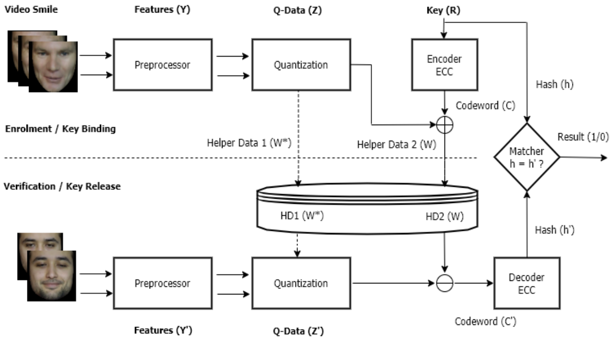

Fig. 1. A System model

At the stage of *Qantization*, the real values *Y* converted into their quantized versions *Z* producing also deviations *W** from the centers of the quantization intervals used as Helper Data 1 (HD1). In the *Encoder* block, the user's password or Key *R* is encoded with one or more ECC codewords, depending on the required password strength and required FRR, which will be discussed below. Further, for a biometric authentication purposes the bit representation of the resulting codeword is added modulo 2 to the quantized and encoded version *C* of smile imprint *Z*. The result is a sequence *W*, which serves as Helper Data 2 (HD2). Further, we will consider the authentication scenario, although this BS can also implement the identification process.

The proposed system works in two operating modes (see Figure 2). In the first mode, the user is registered and his smile imprint is obtained from SAE and with the use of HD1 and HD2 linked to the secret Key *R*, the *h*-hash of which is calculated and stored in the Biometric Database (BD). During verification, the reverse process of decoupling the “auxiliary” data HD1, HD2, decoding *C’* and comparing two hashes *h* and *h’* is performed by a *Matcher*.

Thus, the user smile imprint captured from his video can serve as the biometric key to organize the access to different external digital services. At the same time, the proposed system implements the JW scheme with the use of two helper data HD1, HD2 sets that can be stored in the public domain. To represent information in HD1 and HD2, various methods can be used up to encryption, which is determined by the complexity and required speed of the implemented biometric system. System parameters such as the length of the cryptographic key, the type and characteristics of the ECC must be determined by

Pattern Recognition and Information Processing (PRIP'2021) : Proceedings of the 15th International Conference, 21–24 Sept. 2021, Minsk, Belarus. – Minsk : UIIP NASB, 2021. – 246 p. – ISBN 978-985-7198-07-8.

the quality of biometric data, the number of users, the resistance to external attacks, etc.

IV. EXPERIMENTAL RESULTS

A series of experiments were performed with SAE to get good compact biometric features. To reduce time spent, in these experiments the subsets of 40 subjects randomly selected from the entire UvA-NEMO Database were used, reproducing a posed smile. Then normalized grayscale images from corresponding video of 112x112 pixels in size, scaled to 50%, creating a vector length of the input layer of 6272 elements were used for unsupervised learning of SAE. Image scaling made it possible to reduce the time spent on data processing while maintaining sufficient differentiation of users' smiles when using the selected autoencoder structure. Combinations of the second and third SAE layer dimensions had values 255/63, 127/63, 127/31, and 63/31. The selected values were determined by the length of the applied RS codes, as well as the chosen dimension of equidistant quantization.

To perform quantization and encoding, the following structures were chosen 127/63 and 63/31. The selected values were determined by the length of the applied RS codes, and the chosen dimension of equidistant quantization.

To evaluate the quality of training, on the basis of latent layer data *y*, such values as FRR, FAR, GAR, ERR and DI were calculated, the values of MSE for controlling the intra-class and inter-class distance distribution and the ROC-characteristic were monitored. The results after unsupervised learning and then supervised tuning of SAE with parameters 127/63 for 40 users, in the form of histograms, are shown in Fig. 2. It can be seen from the figure that due to the fine-tuning procedure, the interclass distributions expanded significantly relative to each other.

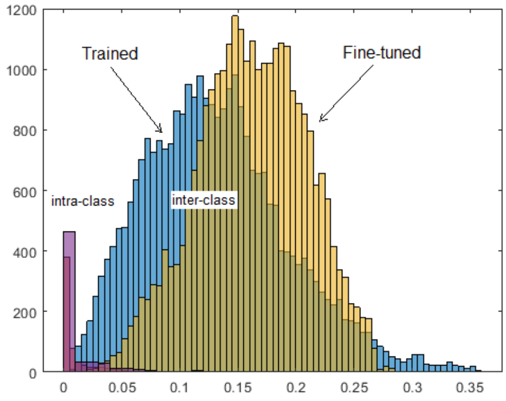

Fig. 2. Learning results of the SAE with parameters 127/63

Then several experiments were carried out to train SAE for 5 different groups of 40 users each, randomly selected from the UvA-NEMO database. For this compiled dataset the encoding-decoding procedures were modeled using the above-mentioned ECC. For processed 200 subjects from the dataset, the real-valued data with lengths of 63 and 31 elements, obtained from SAE and equidistantly quantized, were encoded with non-binary RS codes concatenated with linear codes.

Schemes of ECC used in modeling included: 1). RS code (63,15) concatenated with linear code (6,3,1) and separately with repetition (REP) codes (3,1,1); 2). RS code (31,9) concatenated with REP codes (3,1,1) and (5,2,2) separately. The conversion of symbols to bits was done on the basis of Gray's code representation [10]. For example, using the RS code RS (63,15) with Linear (6,3,1), we can obtain the FRR=1.0% and get the user key *R* with the length of 135 bits from the SAE data extracted from 3 video frames. And if you apply the RS (31,9) code together with REP (3,1,1), then the FRR will decrease to 0.5% for a key *R* with the length of 120 bits, processing 8 frames of user video. Thus, there is a trade-off between a decrease in FRR and an increase in the number of processed video frames. The experimental results showed that the efficiency of the proposed BS based on a neural network and non-binary codes is much higher compared to [11].

The hard-decision decoding has been applied for both type of ECC. An attempt to use the error-reducing technique [12] based on the properties of HD1 together with ECC failed to achieve adequate improvement due to the rather large variance of quantized data. However, in subsequent studies we are considering the possibility of using all types of HD and applying soft decoding of RS codes.

V. CONCLUSIONS

In this article, we examined the principles of implementing the Biometric System based on the use of facial smile imprints. Video frames from the main smile phases (onset, apex and offset) were used to obtain biometric data, on the basis of SAE, which consisted of two inner layers, and was trained with the use of smile videos of 400 subjects, taken randomly from UvA-NEMO database. The real data of the output layer was quantized and encoded by the concatenated ECC based on RS codes (63,15) and (31,9) with a redundancy of more than 70%, which affects the entropy loss or leakage rate [12]. The simulation experiments to assess the system performance showed the possibility of achieving FRR values of no more than 0.5% and 1% for crypto keys

Pattern Recognition and Information Processing (PRIP'2021) : Proceedings of the 15th International Conference, 21–24 Sept. 2021, Minsk, Belarus. – Minsk :
UIIP NASB, 2021. – 246 p. – ISBN 978-985-7198-07-8.

of size 120-135 bits for biometric feature data dimensions of 31, 63 elements.

In our setup, we have used rather simple neural network (NN) structures and concatenated ECC based on non-binary RS codes which made it possible for us to obtain the FRR of less than 1% for the proposed biometric cryptosystem.

The direction of further work can be both the study of other NN structures to obtain deep features of a smile facial imprint, and the improvement of the technique of using ECC based on soft-decision decoding and side information from quantized data.

ACKNOWLEDGMENTS

This material is based upon work partially supported by the COST Action CA16101. The authors would like to thank Dr. Hamdi Dibeklioğlu for providing access to the data base UvA-NEMO.

REFERENCES

[1] Smileid. New Standard for Face Biometrics. Available from: https://www.electronicid.eu/en/solutions/smileid, last accessed 2021/05/05.

[2] S. Cook, Selfie banking: is it a reality? Biometric Technology Today, 9–11, 2017.

[3] M. Taskiran, et al. Face Recognition Using Dynamic Features Extracted from Smile Videos. In: IEEE International Symposium on Innovations in Intelligent Systems and Applications (INISTA), Sofia, 3-5 July 2019, pp. 1–6.

[4] H. Dibeklioglu, A. A. Salah, T. Gevers, Are you really smiling at me? Spontaneous versus posed enjoyment smiles. In: European Conference Computer Vision, Florence, 7-13 Oct. 2012, pp. 525–538.

[5] K. Siwek, S. Osowski, Autoencoder Versus PCA in Face Recognition. 18th Intern. Conf. on Computational Problems of Electrical Engineering (CPEE). Kutna Hora, 11-13 September 2017, pp. 1–4.

[6] A. Juels, M. Wattenberg, A fuzzy commitment scheme. ACM Conf. on Computer and Communications Security, Singapore, November 1999, pp. 28–36.

[7] B. Assanovich, AutoEncoders for Denoising and Classication Applications. In: Open Semantic Tech. for Intel. Sys. (OSTIS), 4, 2020, pp. 309-312.

[8] A. J. H. Vinck, Coding Concepts and Reed-Solomon Codes. Institute for Experimental Mathematics, Essen. 2013.

[9] B. Assanovich et all. Recognition of Genuine Smile as a Factor of Happiness and its Application to Measure the Quality of Customer Retail Services. Proc. 14th Intern. Conf. Pattern Recognition and Inf. Proc. (PRIP'2019), Minsk, 21-23 May 2019, pp. 84-89.

[10] K. D. Rao, Channel Coding Techniques for Wireless Communications, Springer India, 2015, 394 p.

[11] B. Assanovich, Yu. Veretilo. Biometric Database Based on HOG Structures and BCH Codes. Proc. Information Techn. and Syst. 2017 (ITS2017). Minsk, 2017, pp. 286-287.

[12] V. Immler, K. Uppund. New Insights to Key Derivation for Tamper-Evident Physical Unclonable Functions. IACR Trans. Cryptogr. Hardw. Embed. Syst. 3, 2019, pp. 30–65.

Pattern Recognition and Information Processing (PRIP'2021) : Proceedings of the 15th International Conference, 21–24 Sept. 2021, Minsk, Belarus. – Minsk : UIIP NASB, 2021. – 246 p. – ISBN 978-985-7198-07-8.