# Cloud Detection in Aerospace Imagery of Environmental Monitoring

Alexander Labokha
ITS Department
Belarusian State University of Informatics and Radioelectronics
Minsk, Belarus
labokha.poit@bsuir.by

Artsiom Shamyna
ITS Department
Belarusian State University of Informatics and Radioelectronics
Minsk, Belarus
shamyna@bsuir.by

Alexei Ardyako
ITS Department
Belarusian State University of Informatics and Radioelectronics
Minsk, Belarus
ardyakon@gmail.com

***Abstract.* In this paper, the problem of cloud detection in the processing of aerospace images is considered for further use of images in monitoring systems for natural complexes. The existing methods of cloudiness recognition on aerospace images are analyzed, an approach based on an admissible range of values is proposed for solving a particular class of problems. An experimental comparison of the algorithms is carried out, and on the basis of the results obtained, the effectiveness of the proposed method is determined. An approach is identified for the most complete removal of cloudiness from the image.**

***Keywords:* cloud detection, aerospace images, monitoring, GIS**

## I. Introduction

Aerospace images can be used in monitoring systems for natural and other complexes. In this case, it is important to improve the regularity of earth remote sensing data updates. The presence of partly or overcast clouds in images can significantly increase the refresh period, so it is important to use informative pixels even from images that contain clouds. One of the possible solutions to this problem is to identify and replace pixels containing clouds with pixels from the previous image with informative data.

The research was carried out as a part of the development of a methodology for determining the level of fire hazard using aerospace images in fire hazardous periods (from March to October) 2019-2020 for the territory of the Volozhin forestry enterprise. During the selected interval (245 days) in 2019, Sentinel-2 satellites took 61 images with a cloud percentage of no more than 20% and 81 images with no more than 40% cloud cover; in 2020, 36 and 64 images were taken, corresponding to a cloud level of no more than 20% and 40%, respectively. So, the average frequency of taking images with an acceptable cloud level is at least 3 days, which confirms the relevance of the task.

## II. Cloud detection algorithms

The Braaten-Cohen-Yang method [1] uses the threshold condition to determine the presence of cloudiness in the considered pixel based on the values in spectral channels B3 and B4 of Sentinel-2 satellite images [2]. The efficiency of this fairly simple algorithm on the Hollstein dataset reaches 73% of the classification accuracy. Most of the errors arise from the definition of snow as a cloud and the inability to detect thin clouds.

A Scene Classification Layer (SCL) layer is available for Sentinel-2 satellite images. The Scene Classification Algorithm [3] generates classification maps that include four different classes of clouds (including cirrus) and six different classifications for shadows, cloud shadows, vegetation, soil or deserts, water and snow. The algorithm detects pixels with clouds or cloud shadows and is used for replacing pixels with clouds based on the SCL classification.

The approaches described above are universal, they affect in their changes only pixels that contain clouds, and do not affect other pixels that can potentially contain information useful for research. However, in the case of solving a certain problem, it is possible to sacrifice pixels, that don’t contain cloudiness, and will not be used in the future. Then, for the pixels of interest a range of values can be determined to be enough large in order to guarantee the extremely low probability of the value of the pixels of interest to fit in range. The pixels with values outside the specified range are determined as potentially containing clouds and they can be replaced with pixels from the previous image.

The interest of this study is an information about forest vegetation, the analysis of the spectral channel

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.

B2 of various images with cloud percentages from close to 0% to 50% and higher was carried out. Information on forest vegetation is predominantly in the range from 200 to 400, while pixels containing cloudiness generally have values greater than 800 and 1000. Pixel values that are greater than 800 and 1000 also corresponded to agricultural areas and settlements. This confirms the previously presented hypothesis and suggests that replacing pixels with a value greater than 1000 (or 800) in spectral channel B2 will not affect pixels with useful information about forest areas, but will replace most of the pixels "polluted" by clouds.

It should be noted that such an approach will not allow identifying all types of cloudiness, and does not allow taking into account the presence of cloud shadows in the images. Since each method has both strengths and weaknesses, it is likely that sequential application of several methods for replacing cloudy pixels will reveal the largest number of cloudy pixels.

## III. EXPERIMENTAL COMPARISON OF ALGORITHMS

When conducting an experimental study, the effectiveness of the described methods was analyzed both separately and as a result of joint application for the same image. For the best assessment of the methods, an image of the Volozhin forestry enterprise from 06/17/2020 was selected, containing various types of clouds, as well as cloud shadows; a fragment of its spectral channel B2 is shown in Fig. 1, a. As the previous donor image contains a low percentage of cloud cover, the image from 06/17/2020 was used; a fragment of the spectral channel B2 in Fig. 1, b. The result of applying the Braaten-Cohen-Young method is shown in Fig. 1, c; method based on the classification of scenes Sentinel-2 – in Fig. 1d; method for analyzing channel B2 values with a range of values from 0 to 1000 – in Fig. 1, e. The result of sequential application of these three methods is shown in Fig. 1, f. For a representative visual comparison, all images are displayed in a gradient from black to white in the range from 0 to 2000.

## IV. RESULTS AND CONCLUSIONS

Visual analysis of the obtained images allows us to make the following conclusions:

1) The Braaten-Cohen-Young method is better than other methods to recognize transparent and translucent clouds, as well as cloud shadows. At the same time, this method is much worse than the others in processing continuous clouds, without replacing individual pixels within the cloud.

2) The Sentinel-2 scene classification method is well suited for replacing pixels within overcast clouds. This method poorly handles transparent and semi-transparent edges of clouds, replacing only part of the pixels in the center of such clouds. The algorithm does not detect pixels covered with haze or shadowed by clouds.

3) The method based on the analysis of the values of the spectral channel B2 with an allowable range of values from 0 to 1000 according to the results of the work is close to the method based on the classification of scenes, but it is better for processing semitransparent clouds and the boundaries of continuous clouds. The handling of transparent clouds and cloud shadows is also low.

4) The combined use of the three considered methods allows you to use the advantages of each of the methods and mutually reduce their disadvantages. The resulting image contains a minimum of cloudiness compared to the results of using the methods separately.

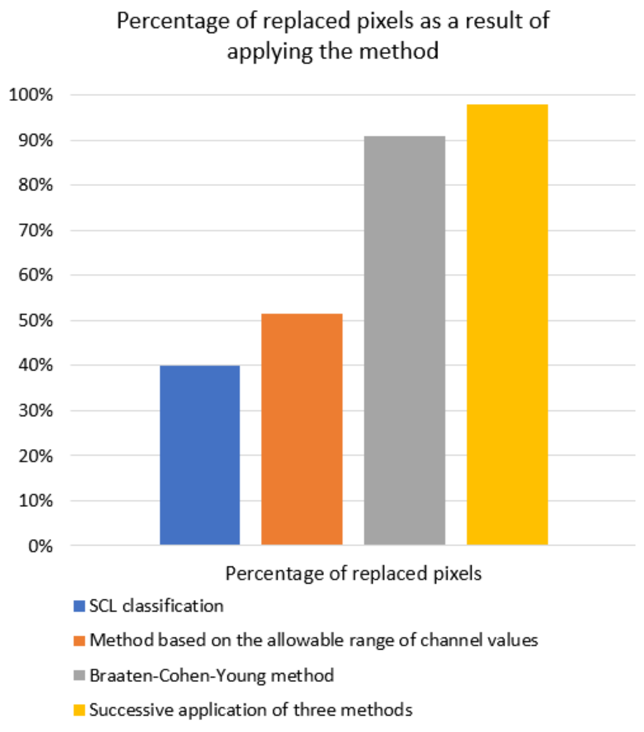


Fig. 2. Histogram of the percentage of replaced pixels

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.

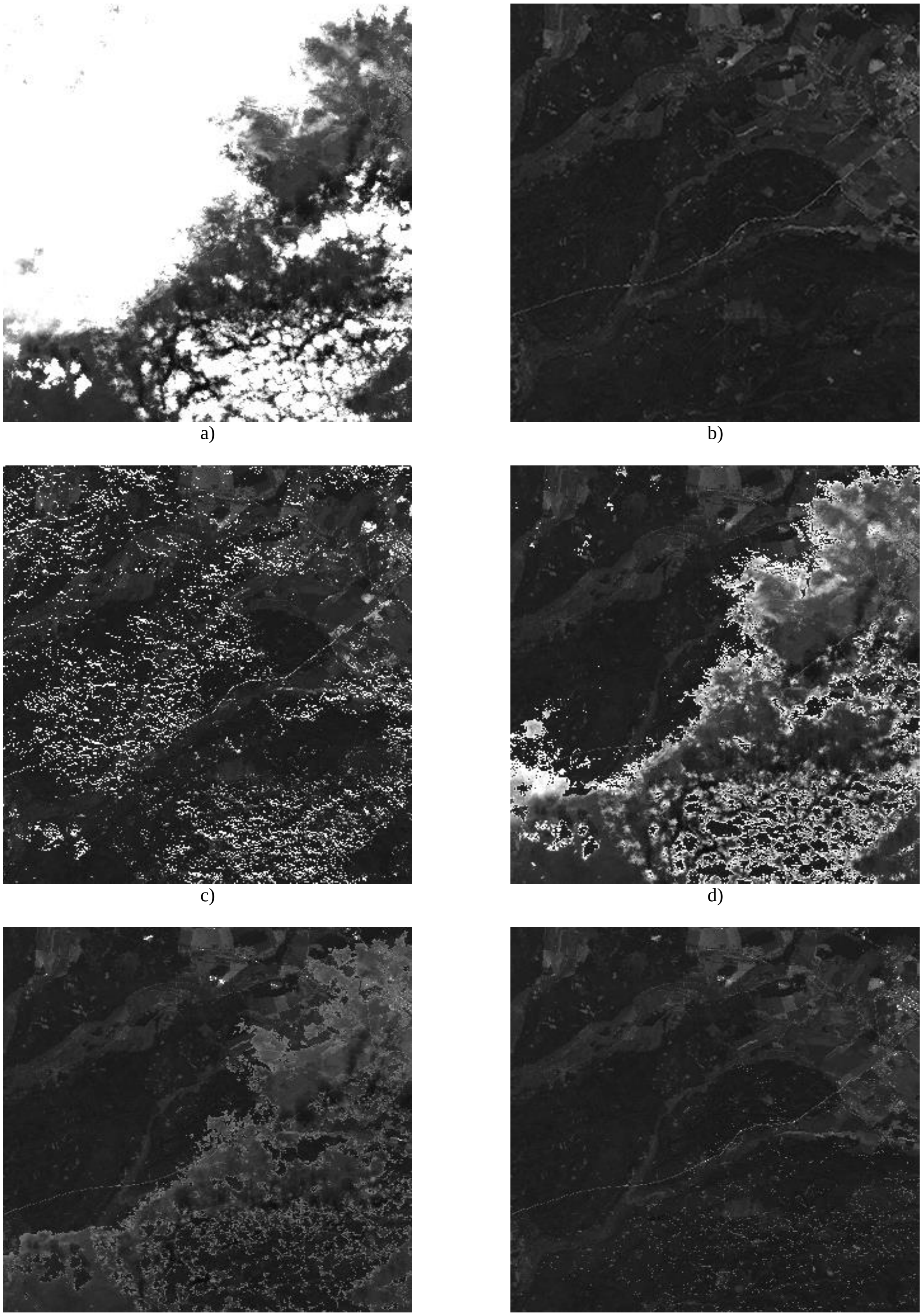


Fig. 1. Spectral channel B2: a) image for June 17, 2020, without processing; b) image for 06/10/2020, without processing; c) image for 06/17/2020, cloud replacement using the Braaten-Cohen-Young method; d) image for June 17, 2020, cloud cover change according to SCL classification; e) image for 06/17/2020, cloud cover replacement according to the allowable range of values from 0 to 1000; f) image for 06/17/2020, replacement of clouds by successive application of three methods

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.

A numerical assessment of the effectiveness of each of the methods is the percentage of detected pixels containing clouds of the total number of pixels belonging to forest areas, it is shown in Figure 2. For a selected day with a high percentage of cloudiness, the method based on the classification of Sentinel-2 scenes marked about 40% of pixels as containing clouds. The proposed approach based on the range of values noted by a quarter more pixels (more than 50%), while the spatial distribution of the detected cloudiness correlates with the results of the SCL method, expanding the selected areas (Fig. 3, a, b). The Braaten-Cohen-Yang method revealed almost twice number of pixels (about 90%) containing clouds compared to other methods. However, upon visual analysis (Fig. 3, c), it can be seen that this method did not reveal individual pixels within the areas of continuous clouds, which were replaced by other methods. Therefore, the combined application of all three methods for the test image revealed the largest number of pixels with cloudiness (98% of the total number), making it possible to combine the advantages of different methods.

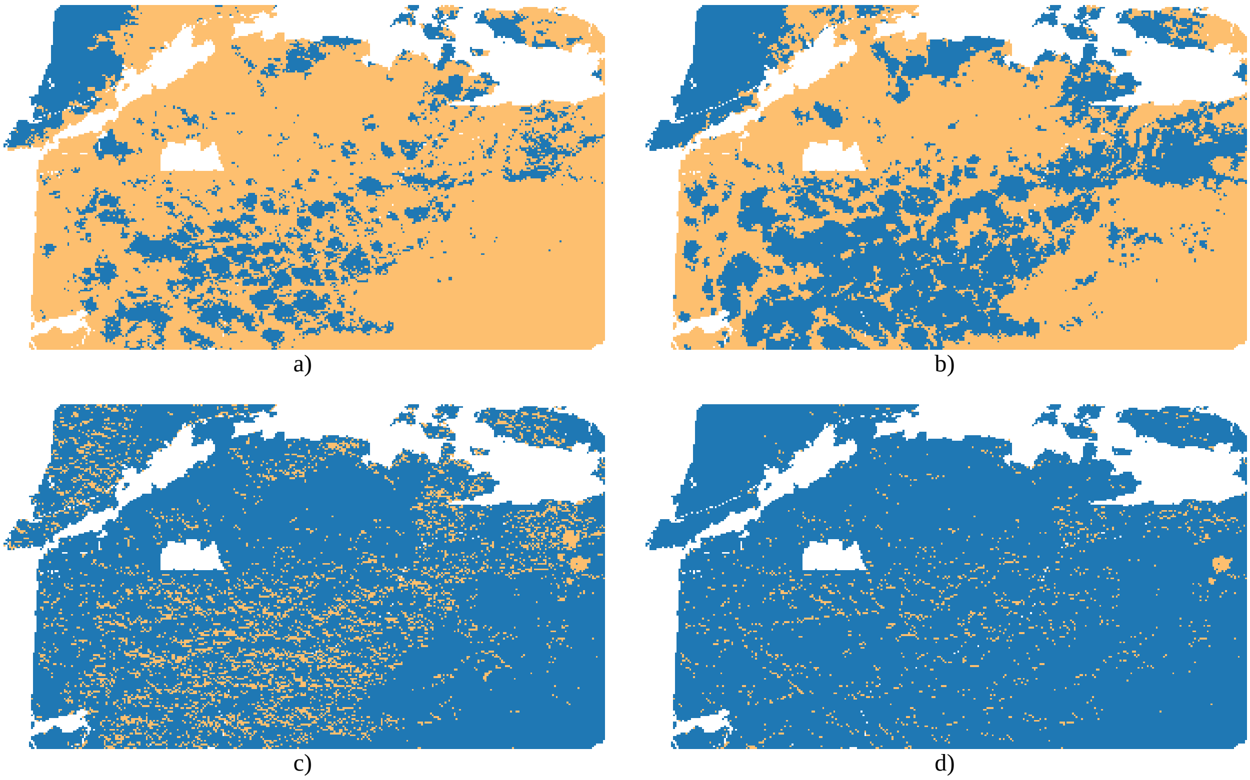


Fig. 3. Map of forest areas (blue color – pixels marked as containing clouds; orange color - pixels marked as not containing clouds; white color - pixels outside forest areas) as a result of cloudiness replacement: a) according to SCL classification; b) by the allowable range of values from 0 to 1000; c) by the Braaten-Cohen-Young method; d) sequential application of three methods

The research clearly demonstrates that the sequential application of various methods for identifying pixels containing clouds is most effective. At the same time, in the case of solving a highly specialized problem, filtering areas based on the allowable range of values is also a reason-able approach and allows you to identify a larger number of pixels containing clouds.

## References

[1] Braaten, Justin, W. B. Cohen and Zhiqiang Yang. "Automated cloud and cloud shadow identification in Landsat MSS imagery for temperate ecosystems." Remote Sensing of Environment, 2015, pp. 128–138.

[2] Braaten-Cohen-Yang cloud detector | Sentinel-Hub custom scripts [Online].Available:https://custom-scripts.sentinel-hub.com/sentinel-2/cby_cloud_detection.

[3] Level-2A Algorithm - Sentinel-2 MSI Technical Guide - Sentinel Online [Online]. Aviabvle: https://sentinel.esa.int/web/sentinel/technical-guides/sentinel-2-msi/level-2a/algorithm.

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.