

The Complex Method to the Solution of the Problem of Polychromatic Images Segmentation

V.V. Bryzgalin, E.A. Danilov, I.I. Salnikov

Abstract: *The article touches upon the problem of polychromatic images segmentation. It is a very urgent issue, because today the sphere of electronic mapping has rapidly been evolving. The research aims at developing a technique of adaptive segmentation and implementing it as a program for a PC. Aerial snapshots taking from the countryside are treated as objects of our research. The developed method combines background areas detection, edge detection and dimensions selection. The experiment has been carried out. It proves the efficiency of the developed technique.*

Keywords: *segmentation, polychromatic image, local area, Sobel operator for edge detection, areas edge detection, grayscale image, aerial photography, electronic map, brightness gradient, images recognition*

I. INTRODUCTION

During recent decades, there has been a rapid development of information technology and the element base of computing devices. In this regard, it becomes possible to implement the tasks associated with pattern recognition and artificial intelligence. One example of such tasks is the recognition of objects on a terrain map represented by a series of photographs from an unmanned aerial vehicle.

Photos can be presented in various color models (palettes):

1) RGB (Red, Green, Blue) - additive color model. Each pixel of the image corresponds to the brightness values of red, green and blue. The model is called additive, since the colors are obtained by adding to black. The choice of primary colors is determined by the physiology features of color perception by human eye retina. The RGB palette is used in programming and computer graphics.

2) HSV (Hue, Saturation, Value) - the color characteristics of each image pixel are set by 3 parameters: color, saturation, value. The HSV palette is based on a person's intuitive perception of color reproduction, and is therefore often used in computer graphics programs. The HSV model is a non-linear transformation of the RGB model.

3) CMYK (Cyan, Magenta, Yellow, Key color) - the color model used in printing to print color images. Each pixel is formed by 4 colors: cyan, magenta, yellow, black. The CMYK model is subtractive (subtracted), since the primary colors of the RGB palette are subtracted from white. This approach is convenient as it is more convenient for paper and other printed materials to consider the amount of light reflected from the surface than absorbed by the surface.

Revised Manuscript Received on September 22, 2019.

* Correspondence Author

V.V. Bryzgalin, post-graduate student, PenzSTU, Penza, Russia,

E.A. Danilov, PhD in technical sciences; PenzSTU, Penza, Russia,

I.I. Salnikov, Doctor of technical science, professor, Head of the "Computing machines & Systems" Department, PenzSTU, Penza

One of the most important steps in pattern recognition is segmentation. Segmentation as the method of digital image processing is the division of the raster into spatially related groups of elements [1]. In [2], segments are also called local regions (LR). The simplest segmentation option is possible if an original image is binary. It is stated in [3]. This method of region localization has the following disadvantages:

- 1) it is impossible to determine the coordinates of the segment maxima in a binary image.
- 2) a binary image does not always convey the outlines of real-world objects accurately.

Therefore, the most correct segmentation is obtained if the original image is full-gradation.

The result of image segmentation is significantly influenced by the criteria of spatial connectivity of pixels. Connectivity is the number and location of neighboring pixels, the brightness of which affects the belonging of a pixel to a particular segment. In [1], they state that 6-connected neighborhood system with diagonal symmetry is most suitable (Figure 1, a). In [2], they proved that four-connected format in the form of a pixel SK (i, j-1) adjacent to the current pixel SK (i, j) for the current i-line and the pixels SK (i-1, j-1), SK (i-1, j), SK (i-1, j + 1) for the previous (i-1) line (Figure 1, б). There is also 8-connected neighborhood system (Figure 1, в).

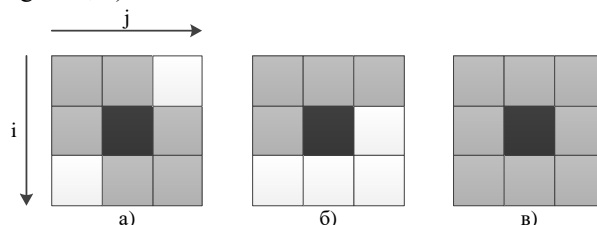


Figure 1 - Variants of spatial connectivity for neighboring elements

This study uses 8-connected neighborhood system due to the fact that it provides the greatest union of local areas.

II. PROBLEM FORMULATION

The research has the following objectives:

- 1) To develop a segmentation method that implements:
 - determination of pixel belonging to segments;
 - definition of boundaries between objects, between an object and a background;
 - coordinate determination of local region maxima;
 - adaptive selection of local area allocation parameters depending on the brightness characteristics of an image.

2) Analysis of the local area parameters of a segmented image for the method operation correctness. To conduct the experiment, the software tool in C ++ was designed. The image of the countryside from an unmanned aerial vehicle was used as a test (Figure 2).



Figure 2 – Test image

During the study, the following assumptions were made:

- 1) The studied images must be bitmap, converted to the Grayscale color model. The Grayscale model means that each pixel in the image has one brightness value.
- 2) The objects extended in space (rivers, roads) can be divided into component parts upon recognition. In the framework of this study, such a separation will not be considered a segmentation error, since often the brightness data of different sections of the same object differ due to the peculiarities of the lighting conditions (Figure 3). For example, on Figure 3, the area 1 and 2 belong to the road, but their brightness is significantly different. Accordingly, certain restrictions are imposed on the object size.

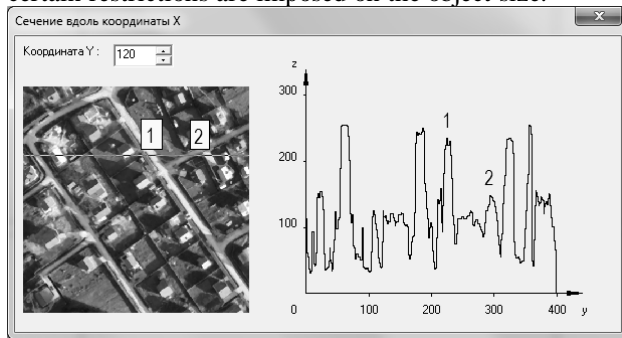


Figure 3 - Test image (left) and brightness section along the image line (right)

- 3) The areas are accepted as the objects, the length and width of which differ slightly.

III. THEORETICAL STUDIES

First of all, it should be decided whether a whole image will be processed by or in parts. From the point of view of the possibility to parallelize operations, it is more convenient to process the portions of an image separately. But on the other hand, the boundaries of such areas can cross local regions, which will lead to incorrect classification of objects on the map. Therefore, as the part of this study, we assume that the image will be processed in its entirety.

As a result of the study, the segmentation method was developed, conventionally called the method of watershed

areas. It got such name as it looks like the division of river basins. Initially, the boundary between the objects is determined, and then the contours of the objects are formed. Segmentation by the watershed method consists of the following steps:

1) Original image smoothing

One of the goals of the study is to determine the coordinates of local region maxima. Having analyzed the brightness section along the line of the test image (Figure 3), we conclude the following: there are many false maxima on the test image that can be mistaken for objects. Therefore, it is recommended to smooth an image before processing by median filtering with the window width of 3x3, 5x5 or 7x7 pixels.

Median filtering is a non-linear signal processing method developed by Tukey. The median filter suppresses scattered impulse noise most effectively [4]. The median filter is a sliding two-dimensional window of N x N by size, covering an odd number of image elements. The central element is replaced by the median of all image elements in the window. The median of the discrete sequence $a_1, a_2, \dots, a_{N \times N}$ for odd N is that the element for which there are (N-1)/2 of elements smaller or equal in magnitude, and (N-1)/2 of elements larger or equal to it by size (1). The pixel brightness matrix of the raster image S_k (1 ... width, 1 ... height) acts as the source data for filtering.

$$\begin{cases} a_{N \times N} = \{a_1, a_2, \dots, a_{N \times N}\}, \\ N \in \{3, 5, 7, \dots\}, \\ a_1 < a_2 < \dots < a_{N \times N}, \\ S_k(i, j) = a_{(N \times N + 1) / 2}, \end{cases} \quad (1)$$

where N is the width (height) of the filter window, $S_k(i, j)$ - the brightness of the pixel with the coordinates (i, j).

The median filtering of images should be considered as the heuristic method in order to suppress noise, therefore, the sizes of the smoothing window N x N should be selected for a specific image. Filtering with the window of 7x7 pixels was used for the test image (Figure 4).

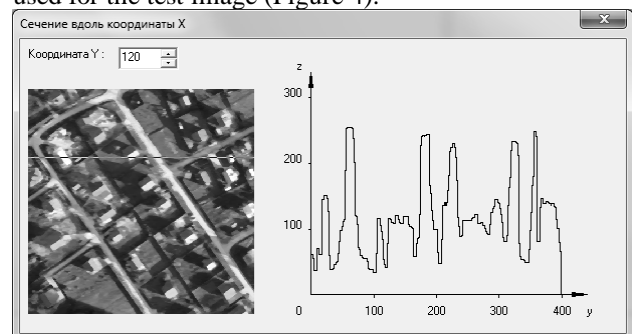


Figure 4 - Test image after smoothing (left) and brightness section along the line of the smoothed image (right)

2) Background area selection

As a rule, the background brightness is less than the brightness of objects, therefore, the minimum brightness of the useful signal is set as a segmentation parameter. At this stage, NS matrix (1 ... width, 1 ... height) is created by the image size to keep the

segment number of each pixel and is initialized with nonzero values. For the pixels with the coordinates (i, j) and the brightness less than the minimum brightness of the useful signal, $NS(i, j) = 0$ (2).

$$NS(i, j) = \begin{cases} 0, & \text{если } S_K(i, j) < S_{MIN\text{ ПС}}, \\ \neq 0, & \text{если } S_K(i, j) \geq S_{MIN\text{ ПС}}, \end{cases} \quad (2)$$

где $S_K(i, j)$ – яркость пикселя с координатами (i, j), where $S_K(i, j)$ is the brightness of the pixel with the coordinates (i, j),

$S_{MIN\text{ ПС}}$ - the minimum brightness of the useful signal.

Let us consider the method of automatic determination of $S_{MIN\text{ ПС}}$ using the example of a test image after smoothing (Figure 4). Let us analyze the histogram of the brightness distribution of the selected image (Figure 5).

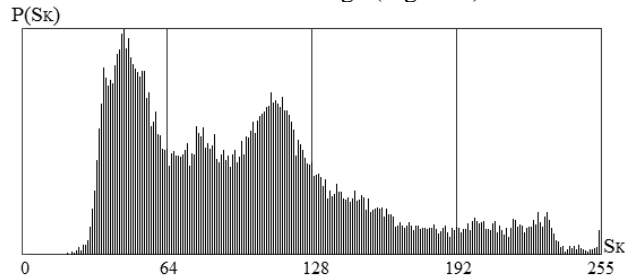


Figure 5 - Histogram of the test image brightness distribution after smoothing

On figure 5, S_K is the brightness level of pixels, $P(S_K)$ is the probability of finding a pixel with S_K brightness on the image. For 8-bit grayscale image $S_K \in \{0, 1, \dots, 255\}$. The left part of the histogram with low brightness refers to the areas of the background, and the right - to the areas of the useful signal. Accordingly, the most accurate value of the $S_{MIN\text{ ПС}}$ is where the probability $P(S_K)$ is minimal and lies between the brightness of the objects and the background. In this regard, it is necessary to limit the studied area of the histogram to search for $S_{MIN\text{ ПС}}$. Indeed, the minimum value of $P(S_K)$ in the histogram areas $S_K = \{0 \dots 30\}$ and $\{230 \dots 255\}$ may indicate a known incorrect $S_{MIN\text{ ПС}}$ value. Search restriction criteria for $S_{MIN\text{ ПС}}$ are given in (3).

$$\begin{cases} S_{MIN\text{ ПС}} \in \{S_{K1} \dots S_{K2}\}, \\ \sum_{S_K=0}^{S_{K1}} P(S_K) = \frac{1}{2}, \\ S_{K2} = S_{MAX} - S_{ГАР}, \end{cases} \quad (3)$$

where S_{K1} and S_{K2} are the boundaries of the search $S_{MIN\text{ ПС}}$, S_{MAX} - the maximum image brightness value, $S_{ГАР}$ - the guaranteed interval of object brightness.

The guaranteed interval of object brightness determines which interval of brightness is not related to the background on the histogram for sure. The value of $S_{ГАР}$ is selected empirically, in the framework of this study, we take it equal to 10% of the total number of brightness levels, that is, equal to 25. For the test image $S_{K1} = 93$, $S_{K2} = 230$, $S_{MIN\text{ ПС}} = 187$ (Figure 6). The test image after highlighting the background areas is shown on Figure 7.

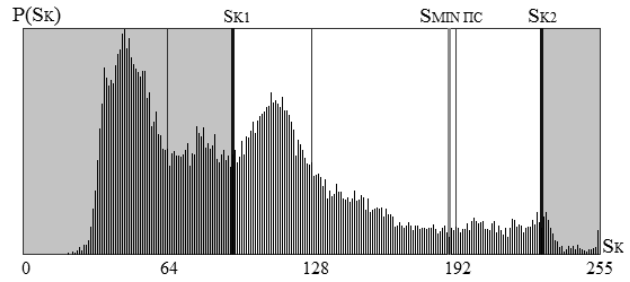


Figure 6 - Determination of the minimum brightness of the useful signal

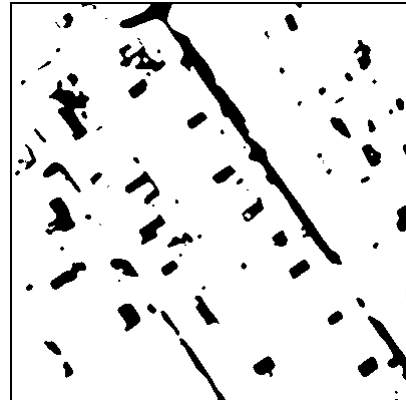


Figure 7 - Test image after background area highlighting

3) Search for the boundaries of the watershed areas

By the boundaries of the watershed areas, we mean the intersection areas of 2 or more LOs. The main objective of this stage is to determine accurately the boundaries of each LO. Edge selection operators may help in this. There are several varieties of operators to distinguish a region: Roberts, Sobel, Kenny and others. More details on edge selection operators can be found in [1,4,5].

In computer vision, the Roberts cross operator is one of the earliest algorithms for border detection. It calculates the sum of value squares of the brightness derivative for diagonally adjacent pixels. For each pixel of the S_K image (1 ... width, 1 ... height), the convolution with two kernels (4) is performed:

$$S_{KP,POB} = \sqrt{S_{KP1}^2 + S_{KP2}^2},$$

$$S_{KP1} = \begin{bmatrix} +1 & 0 \\ 0 & -1 \end{bmatrix} * S_K, \quad (4)$$

$$S_{KP2} = \begin{bmatrix} 0 & +1 \\ -1 & 0 \end{bmatrix} * S_K,$$

where * denotes a two-dimensional convolution operation, S_K - source image,

$S_{KR, ROB}$ - an image processed by the Robert's operator, S_{KP1} and S_{KP2} are two images, each pixel of which contains approximate derivative changes in brightness along the diagonals.

Thus, the value of the Robert's operator $S_{KR, ROB}$ of an image pixel with the coordinates (i, j) can be considered as the module of the brightness vector of the pixel image brightness (i, j). The Robert's operator is used in the cases when a high computational speed is needed. However, Robert's operator loses in comparison with alternatives due to its

considerable sensitivity to noise. It gives lines thinner than other methods of border distinguishing.

The Sobel operator is a discrete differential operator that calculates the approximate value of the brightness gradient of an image. By analogy with the Roberts operator, the result of the Sobel operator application for each image pixel is the module of the brightness gradient vector. The Sobel operator is based on the image convolution with small, separable integer filters in the vertical and horizontal directions, so it is relatively easy to calculate. On the other hand, the approximation of the gradient used by it is not ideally accurate, but sufficient for research purposes. The operator uses 3×3 pixel kernels with which an original image is collapsed to calculate the approximate values of the derivatives horizontally and vertically. Convolution is performed for each pixel, with the exception of pixels located on the edge of the image.

In the course of the study, it was decided to choose the Sobel operator as an optimal option in terms of result quality and the hardware cost of processing. As the result of the Sobel operator, the matrix for selection the edge according to the image size S_{KP} (1 ... width, 1 ... height) is obtained, each cell of which has a gradient vector module. Convolution is performed according to (5).

$$S_{KP} = \sqrt{S_{KP,i}^2 + S_{KP,j}^2},$$

$$S_{KP,i} = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix} * S_K, \quad (5)$$

$$S_{KP,j} = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ -1 & +2 & +1 \end{bmatrix} * S_K,$$

where * denotes two-dimensional convolution operation, S_K - source image,

S_{KR} - an image processed by the Sobel operator,

$S_{KR,i}$ and $S_{KR,j}$ are two images, each pixel of which contains approximate derivatives horizontally and vertically.

The matrix S_{KR} (1 ... width, 1 ... height) is similar to the matrix S_K (1 ... width, 1 ... height), with the only difference being that $S_K \in \{0 \dots 255\}$ для 8-разрядного изображения. is for an 8-bit image. For the edge selection matrix, there is no explicit restriction on the maximum value. After smoothing for the test image (Figure 4), 99.997% of pixels fall into the range of $S_{KP} \in \{0 \dots 499\}$ values within the edge selection matrix.

If for a pixel with the coordinates (i, j) $S_K(i, j)$ is less than the edge selection threshold, then $NS(i, j) = 0$ (6). Let's note that the operation of edge selection complements the processing after the previous parameter determination: the minimum brightness of the useful signal $S_{MIN PS}$.

$$NS(i, j) = \begin{cases} NS(i, j), & \text{если } S_K(i, j) < S_{ПБК}, \\ 0, & \text{если } S_K(i, j) \geq S_{ПБК}, \end{cases} \quad (6)$$

where $S_{ПБК}$ – edge selection threshold.

One of the study goals is to find the values of segmentation parameter automatically. In order to determine the threshold value for the selection of the $S_{ПБК}$ edge, let's analyze the

histogram of the test image brightness distribution after the edge selection (Figure 8).

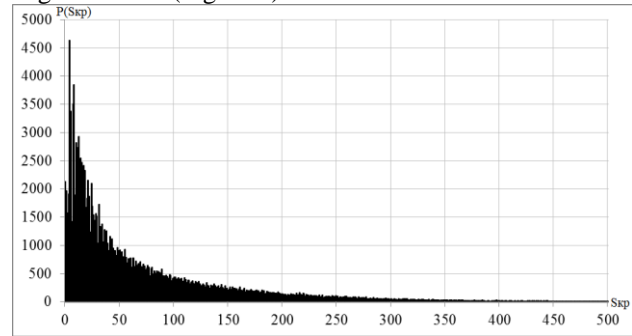


Figure 8 - Image brightness distribution histogram after the edge selection

On Fig. 8, the abscissa axis shows the possible values of the operator for the edge S_{KP} selection; the ordinate axis shows the probability of the of a certain value of $P(S_{KP})$ appearance. The level of 10% of the maximum on the histogram acts as the criterion for $S_{ПБК}$ determination in the framework of this study (Figure 9). The first found value of $P(S_{KP})$, which is less than 10% of the maximum level, is taken as $S_{ПБК}$. The search of $S_{ПБК}$ starts from $P(S_{KP}) = \max$ and is carried out in the direction of S_{KP} increase.

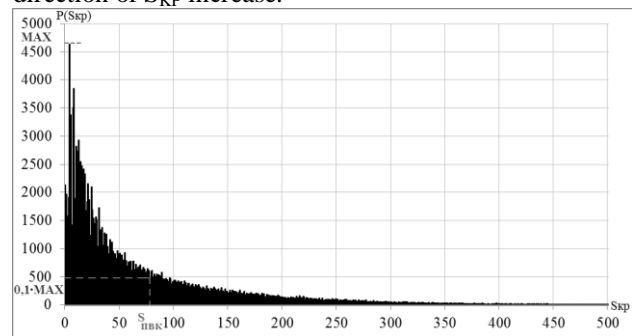


Figure 9 - Finding of $S_{ПБК}$ on the histogram of brightness distribution

For a test image after smoothing, the automatically obtained value $S_{ПБК} = 77$. Figure 10 shows the comparison of the test image after highlighting the areas of the background and after highlighting the edge.

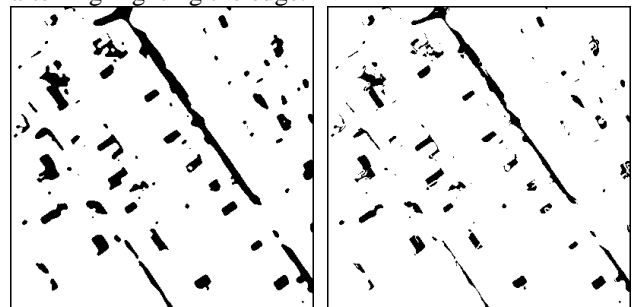


Figure 10 - Test image after highlighting the background areas (left) and determination of the watershed area boundaries (right)

As can be seen from Figure 10, after the operation of edge extraction, noise emissions (plant background) decreased significantly, mainly divided into several parts. The shape of the objects has changed to a lesser extent, since the

brightness of the objects is more uniform.

4) Selection of local areas of the binary image

The matrix NS (1 ... width, 1 ... height) is taken as the basis for the binary image. If a pixel with the coordinates (i, j) has NS (i, j) ≠ 0, then an object is located in this pixel. It remains only to determine its boundaries. To do this, an array of records is formed with the data of all LOs. For each area, the coordinates and brightness of the maximum, the number of pixels are determined. The method is described in more detail by [3]. On the test image, after determination the boundaries of the watershed, some local areas are formed by noise emissions. To eliminate segmentation errors, dimensional selection of regions is used.

5) Dimensional selection of local areas

Initially, the threshold of dimensional selection is set - the minimum number of pixels that should be in the LO. The array of records with the data of all LOs is checked. If the number of pixels of the LO is less than the threshold, then this area is excluded from further consideration. Note that the threshold for dimensional selection is determined empirically, since it depends on the characteristics of a particular image.

6) Determination of local area boundaries

In the framework of this study, it is assumed that the adjacent boundary of the watershed can belong to a certain LO, but the region beyond the boundary of the watershed cannot. Suppose that the border between objects or between an object and a background is at the points with a maximum and minimum gradient (Figure 11).

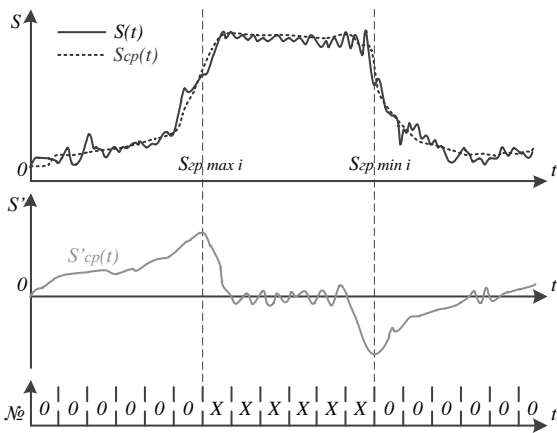


Figure 11 – Determination of LO boundaries in the section of the image row/column

Gradient is a vector quantity characterizing the rate of function change [6]. As the signal S(t), the brightness of the row/column section of the image is used (Figure 11). For function S(t), smoothing is performed by median filtering, thus, a smoothed curve S_{CP}(t) is formed. Based on S_{CP}(t), the graph of the function S'_{CP}(t) derivative is developed, which characterizes the gradient of the function C_{CP}(t) in one of the directions. The gradient for each LO is calculated from the maximum point in 4 directions (horizontal, vertical and two diagonal). The derivative S'_{CP}(t) reaches a maximum or a minimum at the points with a maximum slope of the function graph. The pixels between the maximum and minimum of S'_{CP}(t) are marked as belonging to one of the segments (in figure 11, the segment number is X). For each LO, the brightness level of the boundary is calculated according to

(7). Additional information on image segmentation with the gradient use can be found in [7].

$$\begin{cases} S_{GP} = \frac{1}{n} \cdot \sum_{i=1}^n S_{GPi}, \\ S_{GPi} = \frac{S_{GP\ MIN\ i} + S_{GP\ MAX\ i}}{2}, \\ S_{GP\ MAX\ i} = \max(S'_{CP}(t)), \\ S_{GP\ MIN\ i} = \min(S'_{CP}(t)), \end{cases} \quad (7)$$

where SGR is the brightness level of the LO boundary (minimum LO brightness),
i - the direction of the search for the gradient maximum,
n - the number of directions (in this case, n = 4),
S_{GRi} - brightness level of LO boundary of the i-th direction,
S_{GR MAX i} - brightness level of the maximum gradient of the i-th direction,
S_{GR MIN i} - brightness level of the minimum gradient of the i-th direction.

IV. EXPERIMENTAL STUDIES

The result of the test image segmentation is shown on Figure 12. The “+” symbol marks the maxima of the local regions.

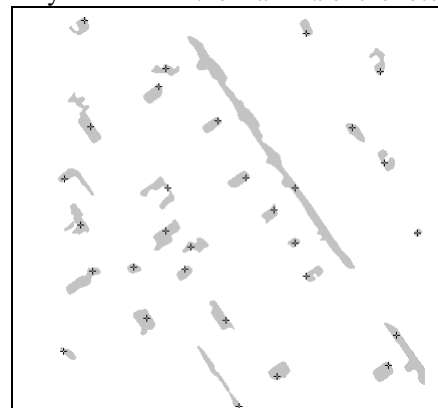


Figure 12 – Watershed segmentation results

To evaluate the effectiveness of a particular image processing method, the statistical brightness parameters of a full-gradation image are analyzed: mathematical expectation, variance, standard deviation (MSD) and signal-to-noise ratio.

The mathematical expectation (average brightness value) is calculated by the formula 8.

$$M\{S_K(i, j)\} = S_{CP} = \frac{1}{N_X \cdot N_Y} \cdot \sum_i \sum_j S_K(i, j), \quad (8)$$

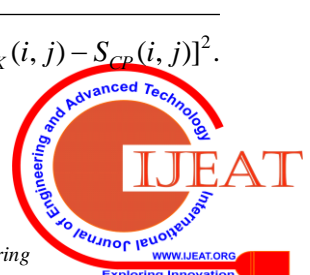
where N_X i- the image width,
N_Y - the image height.

Based on the average brightness value, the variance is calculated:

$$\sigma^2\{S_K(i, j)\} = \frac{1}{N_X \cdot N_Y} \cdot \sum_i \sum_j [S_K(i, j) - S_{CP}(i, j)]^2. \quad (9)$$

MSD is calculated by the formula 10.

$$\sigma\{S_K(i, j)\} = \sqrt{\frac{1}{N_X \cdot N_Y} \cdot \sum_i \sum_j [S_K(i, j) - S_{CP}(i, j)]^2}.$$



Evaluation of the abovementioned parameters is possible for both the color channel and the full-gradation black-and-white image.

The signal-to-noise ratio can be used to assess the distinguishability of objects on images. The signal-to-noise ratio shows how many times the brightness of objects exceeds the brightness of the background.

$$N(M) = \frac{S_{GP.CPE\Delta} - M(S_K)}{\delta(S_K)}, \quad (11)$$

where N(M) is the signal-to-noise ratio of the full-gradation image,

$S_{GR.SRED}$ - the average brightness of the image local area border,

$M(S_K)$ - mathematical expectation of image brightness,

$\delta(S_K)$ - the standard deviation of the image brightness.

The average brightness of the boundary of the image local regions is calculated according to (12).

$$S_{GP.CPE\Delta} = \frac{1}{n} \cdot \sum_{i=1}^n S_{GPi}, \quad (12)$$

where n is the number of segments of the full gradation image,

S_{GRi} - minimum brightness of pixels of the i-th segment.

The average brightness level $M(S_K)$ is considered to be zero, and $S_{GR.SRED}$ is calculated from it. The parameters of the initial and converted images necessary to assess the effectiveness of segmentation results (Figs. 2, 12) are summarized in Table 1.

Table 1 - Watershed segmentation parameters

Image type	Parameter name	Parameter value
Initial	$M(S_K)$ expectation	101,694
	$\delta^2(S_K)$ dispersion	2993,19
	$\delta(S_K)$ standard deviation	54,71
Transformed	$S_{GR.SRED}$ average brightness of LO boundaries	200,1
	N(M) signal to noise ratio	1,799

V. CONCLUSIONS

After the study, they developed the method for segmenting full-gradation images and implemented it as a PC program. All the goals set for the study are fulfilled. The signal-to-noise ratio obtained for the test image $N(M) = 1.799$ is sufficient for the correct recognition of most objects. The adaptability of the method is expressed in the fact that the boundaries of the LO in the image area depend on the characteristics of the objects in this area, such as the coordinates of the maximum brightness and the brightness gradient extrema. This segmentation method will be widely used in terrain image decryption from unmanned aerial vehicles in order to generate an electronic map, as well as in related fields: text recognition, optical control systems. There is a similar solution that implements the watershed method. The Image Processing Toolbox is distributed as the part of MATLAB and implements many computer vision

methods [8]. The disadvantages of the Image Processing Toolbox include the closed source code and the inability to use it separately from MATLAB.

In conclusion, the meaning of “an integrated approach to solve the segmentation problem” should be clarified. If we skip the stage of searching for the watershed boundaries in the considered method, then there will be a lot of erroneously determined LOs as the result of segmentation. On the other hand, the operator of an edge selection without the selection of the background areas first of all will lead to a similar result. In the framework of this study, an “integrated approach” refers to the combination of background highlighting and the search for the boundaries of the watershed areas, which allows you to get the best result.

REFERENCES

1. **Horn B.K.P.** Zrenie robotov (Robot vision). – M.: Mir, 1989. – 487 p.
2. **Sal'nikov I.I.** Rastrovye prostranstvenno-vremennye signaly v sistemah analiza izobrazhenij (Raster space-time signals in image analysis systems). – M.: Fizmatlit, 2009. – 248 p. (in Russian).
3. **Bryzgalin V.V.** Issledovanie metodov segmentacii binarnyh izobrazhenij (Binary image segmentation research). // Voprosy radioelektroniki. Seriya SOIU. no. 2. – M.: CNII «Jelektronika», 2015. – pp. 131-140. (in Russian).
4. **Prjett U.** Cifrovaja obrabotka izobrazhenij (Digital image processing): Per. s angl. – M.: Mir, 1982. – vol. 2 – 480 p., il.
5. **Gonsales R., Vuds R.** Cifrovaja obrabotka izobrazhenij (Digital image processing). – M: Tehnosfera, 2005. - 1072 p.
6. **Dubrovin B.A., Novikov S.P., Fomenko A.T.** Sovremennaja geometrija: Metody i prilozhenija (Modern geometry: Methods and applications). vol. 1. — M.: Nauka. Gl. red. fiz.-mat. lit., 1986. - 760 p. (in Russian).
7. **Sal'nikov I.I.** Analiz prostranstvenno-vremennyh parametrov udalennyh objektov v informacionnyh tehniceskix sistemah (Space-time characteristics analysis of distant objects in informative-technical systems). – M.: FIMATLIT, 2011. – 252 p. (in Russian).
8. **Zhuravel' I.M.** Kratkij kurs teorii obrabotki izobrazhenij (Theory of image processing short course) [Jelektronnyj resurs] / Zhuravel' I.M. – Rezhim dostupa: <http://matlab.exponenta.ru/imageprocess/book2/>. – Zagl. s jekrana. (in Russian).