Multi-threshold Object Selection in Remote Sensing Images

Vladimir Volkov
Saint Petersburg Electrotechnical University (LETI),
Saint Petersburg State University of Aerospace Instrumentation
Saint Petersburg, Russia
vladimi-volkov@yandex.ru

Oleg Markelov, Mikhail Bogachev
Saint Petersburg Electrotechnical University (LETI)
Saint Petersburg, Russia
OAMarkelov@etu.ru, rogex@yandex.com

Abstract—**The problem of selection of objects of different nature in digital monochrome images obtained by remote observation systems is considered. Novel schemes of multi-threshold image analysis and object selection of with the threshold adjustment based on the selection results are proposed. Algorithms for multi-threshold selection of objects by their areas as well as other scale-invariant geometric criteria, such as the perimeter and main axis elongation coefficients are proposed and evaluated. Results of standard model images analysis acting as a testbed as well as real-world images obtained by television and infrared remote systems are considered.**

I. Introduction

The problems of detection, selection and localization of objects of different shapes arise in remote sensing systems including radio systems with synthesized antenna arrays (SAR), infrared, and television systems. In this case identification, tracking, and matching objects as well as combining images from heterogeneous sensors, indexing and image recovery are the common steps comprising the overall image analysis algorithm.

The variety and variability of shapes and textures of objects, as well as their intensive non-stationary background make remote sensing imagery processing non-trivial. Areas of interest are often characterized by small signal-to-background noise ratios, registered digital images exhibit low quality, being quantized to a small number of levels and fuzzy boundaries of both natural and artificial structures such as rivers, roads, bridges, buildings that should be successfully separated from the background noise.

In practice, in remote sensing systems the statistical background is very different from the Gaussian one, the intensity distribution is clearly asymmetric, and the tails of the distributions are similar to the lognormal density of normal or mixed (contaminated-normal), hardly identified from a short data sample. The background can also contain elements that are structurally similar to signals. This nature of the background makes most of the known adaptive thresholding methods inapplicable processing, as incorrect setting of thresholds can lead to loss of useful objects at an early stage. Another problem is the low quality of the generated images, spots, blurred borders. In addition, SAR images suffer from serious internal speckle noise [1].

In both theory and practice of object recognition in remote sensing images and computer vision systems, generally two main approaches are being used. The discriminant approach is based on the comparison against a given template, including correlation- and spectral analysis based methods. Their common drawbacks include weak resistance against possible distortions of individual elements of the objects description.

In contrast, the structural-syntactic approach appears more promising being associated with the analysis of the internal structure of objects of interest and the allocation of their local features. This approach typically includes the stage of transformation or reconstruction of the original image in order to be able to use the selection of objects according to given criteria. In most scenarios, the reconstruction stage aims at the image segmentation, i.e. identification of areas homogeneous according to given criteria. At this stage, the connectivity and homogeneity properties of the regions are used. At the next stage, the regions are selected according to the specified features of the objects of interest. The difficulties of using this approach are associated with the need to describe the diversity of objects by a finite system of rules. In addition, conventional approaches are characterized by a large amount of calculations required, especially when certain geometric transformations to the objects have to be performed.

"Object-oriented image analysis" uses several concepts. Intensity uniformity and color matching are the most commonly used characteristics for image segmentation. Regional methods are mainly based on the assumption that neighboring pixels within the same region have the same value [2].

II. Previous Work

An excellent overview of modern methods of image segmentation is given in [7], where four categories are distinguished, each of which is based on its key element. Such key elements are: a) pixels; b) edges; c) areas; d) others. The first category includes methods of threshold processing and clustering, the second – boundary detectors. The third category includes area growing, watershed method, separation and fusion, level set, and active contours. The remaining fourth category includes the use of wavelets, neural networks, fuzzy sets.

ISSN 2305-7254

Each method has its own possibilities of accounting a priori information about the objects of interest. It is noted that methods based on domain properties, such as fractal net evolution approach – FNEA and methods based on graph theory (Graph methods), dominate the creation of compact object support domains and for acceptable scales. Methods using graphs represented by four basic algorithms: the best merge, a minimum spanning tree (MST), the minimum average cut or minimum mean cut, and normalized cut [7].

There are two ways to form feature areas. One of them ("bottom-up") is based on merging of smaller objects into larger ones, which is used, in particular, the property of homogeneity (it is the best merge and MCT). Another method ("top-down"), on the contrary, considers the original image as an initial single segment, followed by its fragmentation into separate parts (minimum means cut, normalized cut).

These approaches are very constructive for segmentation of images obtained by remote sensing systems (laser locators, radars with synthesized aperture, multi-spectral, hyperspectral, panchromatic, etc.). However, they also have significant limitations. First of all, it is the complexity of computational procedures associated with the solution of optimization problems, and the requirement of high performance of calculators, because often the number of objects to be constructed is very large. In addition, the methods do not have optimality in the choice of the starting points in a sequence of iterations, often resulting in ambiguous decisions depending on changes of initial conditions.

These disadvantages can be somewhat reduced if it is better to use the specifics of the selected objects, training, as well as a combination of different methods to overcome the shortcomings of each of them [7-9]. Ultimately, all the methods under consideration are reduced to organizing image pixels into some multiscale hierarchical structures, which further allow the use of different criteria for selection of objects. The challenge is to make such a structure more transparent and not so difficult to use.

The authors [7] note objective trends in the development of these methods, primarily the development of new theories for more special technologies. This theory can solve the problem of creating a hierarchical structure of objects without using traditional segmentation methods. The key to the new theory may be the use of object selection results in multi-threshold processing, and the application of percolation idea for non-parametric description of object properties.

Multi-thresholding transforms the original monochrome image into a set of binary images (slices). In the case of a sufficiently large number of thresholds, it can be assumed that there is no loss of information with such a transformation. At the same time, binary image processing is easier and faster than multi-level image processing.

Various applications of multi-threshold processing for image segmentation are considered in numerous works, only some of them can be specified [10-14]. Basically, multi-threshold segmentation is based on the histogram properties of the original image. In most cases, the last step is to select the optimal threshold. Properties of objects of interest and the results of their selection are not taken into account. Also, there are no links between different sections of the object at different thresholds.

The expected properties of objects are used in many object selection approaches. Typically, there is an acute lack of information about objects, except for the average size and some assumptions about the area, perimeter, shape and orientation.

The new idea is to choose and set the optimal threshold value based on the results of selection of objects in multi-threshold processing to achieve the best selection of the object having a given size or other geometric parameters. This approach is proposed in the work [3]. Further development of this idea involves the evaluation of a certain geometric parameter of an object in binary images after multi-threshold processing and appropriate selection of objects, and the selection of the optimal threshold value in accordance with the extremum of this parameter. This geometric parameter can be the square of the object or the ratio of the perimeter squared to the square of the entire object [4], [5].

Traditional segmentation schemes use features that stand out from the original image and only indirectly take into account the properties of the objects of interest. In particular, the properties of the histogram of the original image, the properties of edges and contours are widely used. The results of the subsequent selection of objects are practically not used for segmentation. This article describes the method of selection of connected objects based on multi-threshold processing and setting of the threshold according to various geometric criteria and their combinations, based on the approach recently suggested in [3-6].

III. Multi-Threshold Processing

A. *Problem Statement*

Let a monochrome image $I(x, y)$, where I is the intensity and x, y are coordinates of the pixels, is binarized with fixed global threshold T. The result is a binary layer $I_T = \{1, \text{if } I(x, y) \geq T \text{ and } 0, \text{if } I(x, y) < T\}$ in which a subset of units represents the objects of interest (foreground), for example, buildings, structures, vehicles, waterfront, and a subset of the zeros belongs to the background, which is determined by the landscape of observation.

If the pixel intensities of the objects of interest are mainly higher than the background pixel intensities, the generally accepted method of selecting a global threshold is the Otsu method, which operates under rather general conditions. It works based on the analysis of the histogram of the original image and gives the minimum sum of intragroup variances for subsets $\{I(x, y) \geq T\}$ and $\{I(x, y) < T\}$ respectively. It is clear that in practical scenarios the objects of interest are blurred and the background is inhomogeneous and noisy, which in the case of the global threshold leads to overlapping plots of distributions for the specified subsets and unavoidable errors.

Ideally, each object of interest requires its own threshold value, and such local thresholds can be formed by using local (moving) windows, within which the background is considered homogeneous. It is necessary to set controlled or base (associated with the expected dimensions of the object) and the reference (background) area. The reference area is located near the controlled area and is used to form an adaptive threshold [15]. These methods require a priori knowledge of the size of the object of interest. In addition, the use of the background window leads to a loss of resolution on close objects, and to the suppression of one object by neighboring objects that fall within the scope of this window.

Alternative approaches are possible using multi-threshold processing. One of the methods suggests setting a threshold for each category of objects of interest, which is selected by a given criterion [3-6]. In this case, to describe the category of objects, you can use various parameters that reflect, for example, the area size of the object, or its orientation. To analyze multi-scale images, it is more convenient to use invariant characteristics, such as the ratio of the perimeter square to the area, the compression ratio of the ellipse stretching, and other geometric or textural characteristics.

In this case, each binary layer selects objects that satisfy the specified properties, and the binarization threshold for such objects is selected in such a way as to obtain the maximum number of selected objects of this category (or their pixels), taking into account the required preservation of the shape of objects. This process can be automated, resulting in adaptive threshold setting methods.

B. Object Selection by Square

It is believed that the main property that distinguishes the object of interest from the background noise is the connectivity of adjacent pixels in the binary image. In Fig. 1 a model of a monochrome noisy image on a 256x256 grid is presented, in which the signal field (on the left) is summed with Gaussian noise (shown in Fig. 2) in such a way that in the resulting field (on the right of Fig. 1) the objects of interest have a small signal-to-noise ratio d = 1.163 in each signal pixel. It is introduced as the ratio of the shift of the expectation to the mean square value of the noise. The signal field contains rectangular objects of 20x8, 20x16, 20x32, and 20x64 pixels so that the smallest area is 160 pixels.

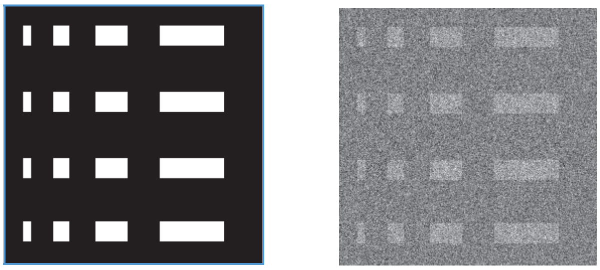


Fig. 1. Test image containing rectangular objects and noisy observation

Let us first consider a purely noise field (Fig.2). Each threshold value displays related objects that contain a different number of pixels. Their number depends on the threshold value [3,6]. Since the shape of the object of interest will be important in the future, small objects will be excluded from consideration, which can significantly reduce the noise level after binarization.

In Fig. 3 the result of single-threshold selection of connected objects of rectangular shape taking into account the removal of small objects is shown. Two types of object shape distortion are noticeable: loss of pixels in the object area and addition of extra pixels along its borders. At high threshold values required for a small number of false objects, the useful objects mostly lose pixels. At small signal-to-noise ratios, useful objects undergo significant boundary deformations that take on a fragmented appearance. This leads to a rather noticeable increase in the perimeter of such connected fragments.

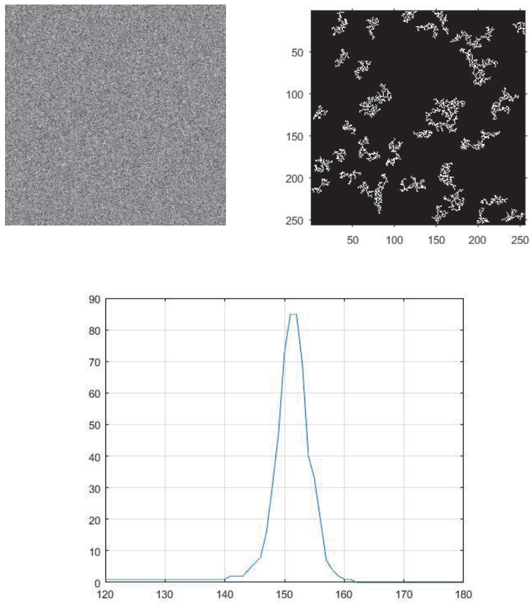


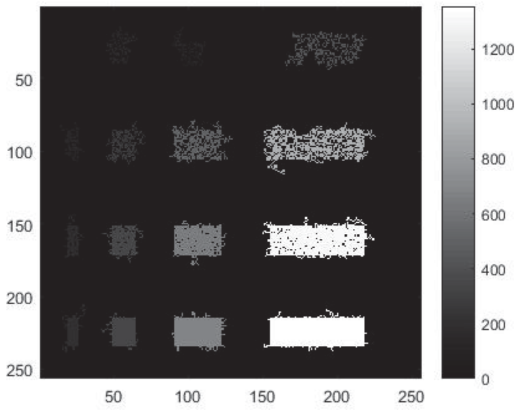
Fig. 2. Noise binarization and the dependence of false objects number from threshold value

The optimal threshold should ensure acceptable preservation of the shape of useful objects. In particular, you can require approximate equality of the number of pixels lost inside the object and the number of pixels "glued" on its border. In this case, the optimal threshold will not correspond to the maximum of the selected objects of the specified area (S_b = 120), but will be slightly shifted towards higher values. In Fig. 3*a* the binarization threshold was T = 115, while the maximum number (N_{obj} = 21) of objects larger than S_b was formed at T = 108 (Fig. 3*b*).

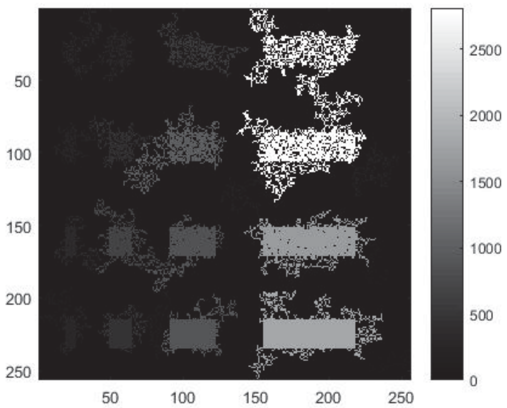
When the threshold decreases (Fig. 3*b*) there is a "gluing" of background pixels along the boundaries of objects, then these processes grow, and then the neighboring objects merge to form conglomerates. In this case, the number of useful objects may decrease. However, false objects appear in the background area, the area of which is comparable to the area of useful objects.

Adaptive selection of objects by area can be implemented by setting a threshold based on the results of multi-threshold processing in such a way as to obtain the maximum number of objects of a given area (or the maximum number of pixels in the selection of such objects), taking into account the requirements for distortion of the shape of objects. In this case, a range of areas of the desired objects of interest can be selected for selection.

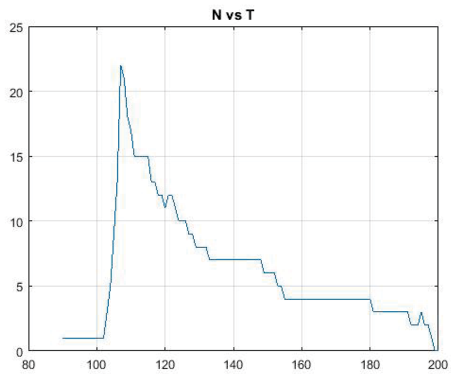
If we assume that the intensity values in the image pixels are mutually independent, and the background and objects of interest are homogeneous, we can calculate the efficiency of detecting the object of interest in a given area S, including n pixels. If the binarization threshold is high enough, then a small number of background pixels that "stick" to the object in the form of fractal tails can be neglected.



a)



b)



c)

Fig. 3 The result of single-threshold selection of rectangular objects and the dependence of the number of selected objects on the threshold value: *a*) T=115; *b*) T=108.The brightness of objects is proportional to their area

Then the problem of detecting the object of interest in noise is solved by registration of k exceeding the threshold of n possible in the region S and comparing the statistics k with the threshold of the account k_T (binary integration method) [15],[16].

In its pure form, the binary integration method can be implemented by summing the number of excesses within the sliding window of specified dimensions, consistent with the size of the detected objects. At each position of the sliding window statistics k is distributed according to the binomial law. The probability of reaching or exceeding the k_T threshold by statistics k is given by the well-known formula [16].

$$P(k \geq k_T) = \sum_{k=m}^{n} C_n^k p^k (1-p)^{n-k} ,$$

where C_n^k – binomial coefficients, and p is the probability of exceeding the threshold in each pixel, this probability is p_0 in the noise background area, and is p_1 in the object area, and it is assumed that $p_1 > p_0$. For sufficiently large n, the binomial distribution can be approximated by Gaussian, and the deflection of the decisive statistics can be introduced as the ratio of the shift of the mathematical expectation to the mean square value of the noise. In binary integration, statistics k has the expectation of $m = np$, and the variance $\sigma^2 = np(1–p)$. Thus, both the expectation and the variance of the decisive statistics change in the area of the object.

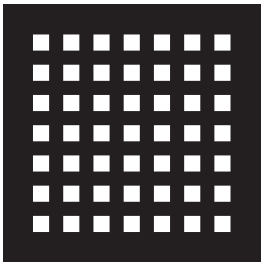
In the case of selection of objects by area, statistics k is no longer subject to binomial distribution, since only connected objects are selected, and their number is significantly less than the number of combinations of n by k. By analogy with the case of binary integration, the probability of reaching or exceeding the k_T threshold by statistics k can be written as

$$P(k \geq k_T) = \sum_{k=k_T}^{n} B_n^k p^k (1-p)^{n-k} , \qquad (1)$$

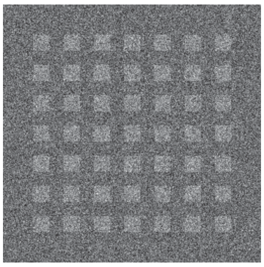
where B_n^k – coefficients whose values determine the number of connected objects consisting of k pixels on the area of n pixels. It was possible to calculate the exact values of these coefficients only for the one-dimensional model and for a small area of objects $n \leq 9$ [6].

The difficulties of the probabilities calculus according to the formula (1) interfere with the determination of the exact threshold value of the account k_T. However, this can be done by adaptation. For the adaptive setting of the threshold, the selection of objects by area is used, taking into account the restrictions on the distortion of the object shape. These distortions are quantified by the number of false pixels appearing around the boundaries of the object and the number of useful pixels disappearing inside the object.

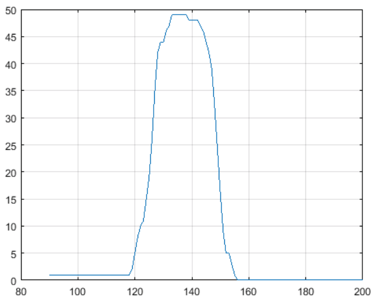
The simulation results are shown in Fig. 4, where 49 square objects of 16x16 pixels size (Fig. 4*a*) are located on a standard Gaussian noise background (Fig. 4*b*). The signal-to-noise ratio (the relative shift of the mathematical expectation) at each pixel is d = 1.163. The dependence of the total number of connected objects on the threshold value is shown on the right (Fig. 4*c*). At selection of objects on the area the acceptable distortion of object borders is reached at the threshold values exceeding $T = 138$ (Fig. 4*d*).



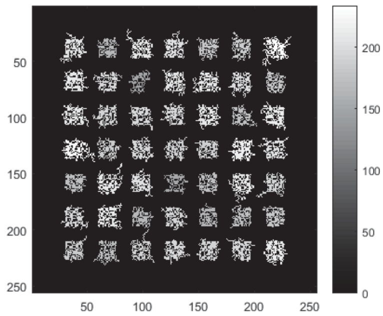
a)



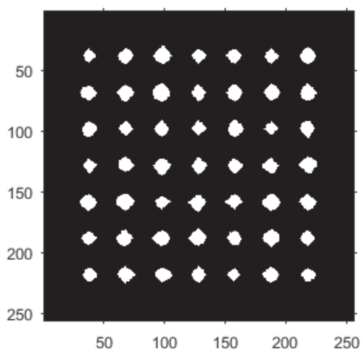
b)



c)



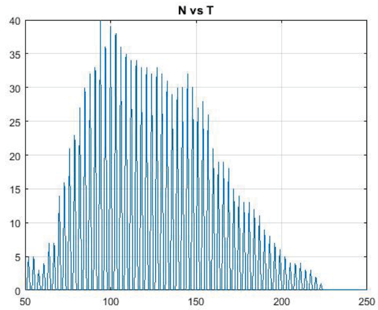
d)



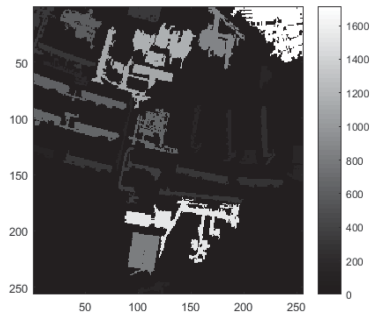
e)

Fig. 4. Modeling of selection by area for square objects (a) on the background of Gaussian noise (b); c) – dependence of the number of selected objects on the threshold value; *d*) – results of selection by area; *e*) – results of object detection by binary integration method.

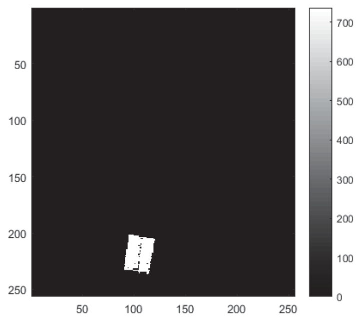
a)



b)



c)



d)

Fig. 5. *a*) – real image; *b*) - dependence of the number of connected objects on the threshold value; *c*) - selection results by area: merging objects at low thresholds; *d*) – selected object.

At lower values of the threshold, the shape of the objects is significantly distorted by fractal noise, which significantly fragments the boundaries. The method of selection of objects by area provides a good selection of the shape of objects even at small signal-to-noise ratios, almost as good as the method of binary integration (Fig. 4*e*). It is clearly seen that the algorithm working by the method of binary integration distorts the shape of objects very much, providing instead the greatest signal-to-noise ratio during accumulation.

Fig. 5 contains results of object selection by area in the image obtained by aerial photography (Fig. 5*a*). Fig. 5*b* shows the dependence of the number of selected connected objects on the threshold value. Image (Fig. 5*c*) obtained by setting the threshold to the maximum number of allocated connected objects (N_{obj} = 35 at T = 101). The scale of intensities reflects the values of the squares of the objects in pixels. With increasing threshold, it is possible to increase the resolution of objects, but less intense objects disappear. If the objects are isolated, then after selection each object is localized (Fig. 5*d*), i.e. the coordinates of its center, as well as other parameters of shape and texture are determined.

The disadvantage of selection by area is the need to set the area parameter in absolute values (in pixels), which is difficult in cases of changing the scale of the image. This method does not work well in the case of non-uniform background, which can give false objects, comparable in area with objects of interest (Fig. 5*c*).

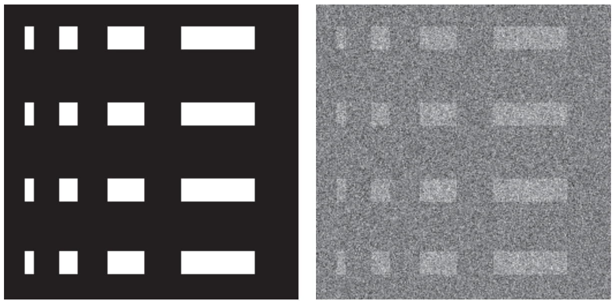
C. Object Selection by Geometrical Invariants

Among the many characteristics of the shape of objects of interest, the most used are those that are invariant to the scale of the image, and the calculation of which is quite simple.

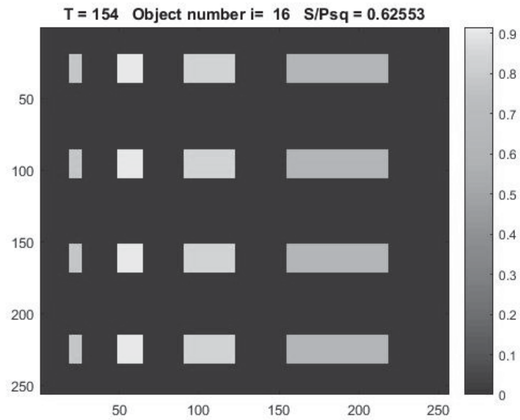
These include such invariants as the ratio of the perimeter square to the area of the object (the perimeter elongation coefficient) and the ratio of the axis lengths of the equivalent ellipse (the ellipticity coefficient).

As is known, the ratio of the perimeter square to the area has the smallest value of 4π for round objects. For convenience, we define the elongation coefficient of the perimeter of the object as $P_S = P^2/4\pi S$, where P is the perimeter of the object, S is its area. Then for the object in the form of a circle, this coefficient is equal to one. Any changes in the shape of the object result in an increase in this factor. The effect of noise leads to fragmentation of the boundaries and their elongation, so that in noise this coefficient also increases for objects of circular shape. Often it is more convenient to operate with an inverse value of $1/P_S$, which varies from zero to one, and has a maximum unit value for the circle.

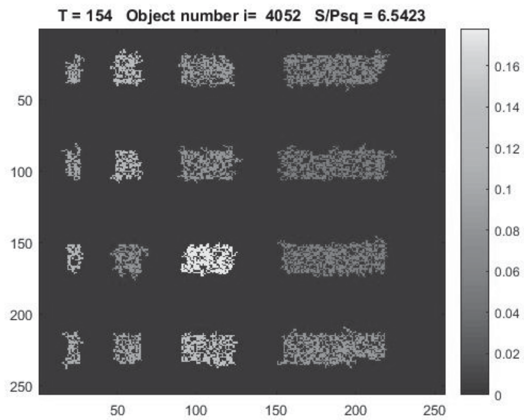
Although formally the minimization of the coefficient P_S is justified for round objects, she is also suitable for the allocation of other regular objects with a coefficient of elongation of the perimeter, close enough to zero, for example, for square objects where this ratio is 1.273. The effect of noise leads to a significant (one-two orders of magnitude) increase in P_S for such objects due to the lengthening of the boundaries, which exceeds minor differences in the coefficients for different regular forms of objects in the absence of noise. Both low and high threshold values lead to an increase in the elongation coefficient of the perimeter of the object under the action of noise. By minimizing this factor, you can choose the best binary layer (and binarization threshold) to represent circular or other regular shape objects (if present in the image) with the least fragmented boundaries.



a) *b*)



c)

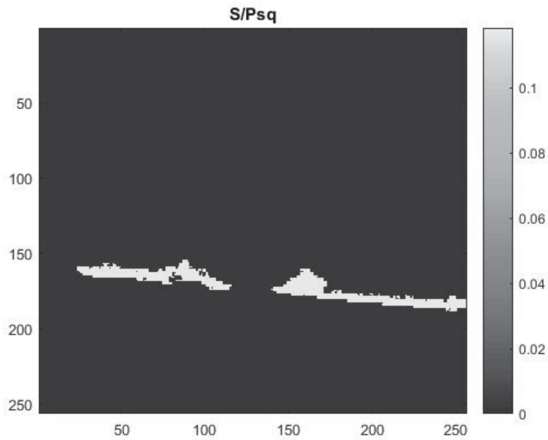


d)

Fig. 6. Selection of objects on the test image by the ratio of elongation of the perimeter. The inverse value of $1/P_S$, which corresponds to the scale on the right, was calculated.

In Fig. 6 results of selection of objects of a rectangular form (with noise) on the inverse value of P_S coefficient are presented. In the absence of noise, $1/P_S$ has the highest value of 0.786 for square objects (the second column in Fig. 6*c*). In the case of a rectangular shape, the objects are lengthened, and this value has smaller values. Effect of noise at low signal-to-noise ratios (Fig. 6*b*) leads to the destruction of the shape of objects (Fig. 6*d*) and to the appearance of fractal processes along the boundaries of objects. As a result, the perimeter of the object grows significantly, and the ratio of the area to the square of the perimeter decreases, which is also facilitated by the disappearance of pixels in the object area. In Fig. 6*d* the maximum value of $1/P_S = 0.178$ has a rectangular object in the third column, although despite the distortion of the form due to noise, square objects have received on average greater values of this value than rectangular.

a)



b)

Fig. 7. Selection of objects on the infrared image by the perimeter elongation coefficient. The inverse value was calculated.

Fig. 7 represents the result of selection of elongated objects according to the criterion $1/P_S$ in the infrared image.

Since the estimation of the perimeter of the object under the action of noise becomes overestimated, it is possible to use instead the estimate of the main axis A of the described ellipse, which reflects well the elongation of the object, while smoothing the fluctuations due to fractality. Consider the geometric invariant to the scale equals to the ratio of the area of the object to the square of the main axis of the describing ellipse. In the case of a circle, the ratio of the square of the diameter to the area is $4/\pi$, so the elongation coefficient of the main axis of the object can be normalized to this value so that for round objects it is equal to one. Thus we define $P_L = A^2\pi/4S$. For a square object, this coefficient is $P_L = \pi/2 = 1.571$. To get maximum values for round and square objects relative to other objects, you can use the inverse of $1/P_L$, which is 0.637 for a square object.

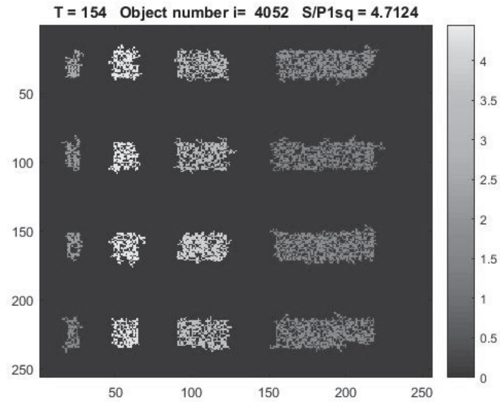
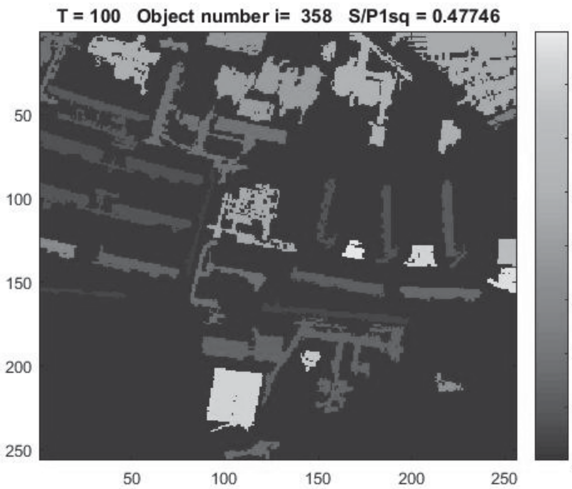


Fig. 8. Selection of objects on the test image by the elongation coefficient of the main axes. The inverse value was calculated.

Fig. 8 shows the results of selection of rectangular objects with a small signal-to-noise ratio at each pixel (d = 1,163) and using the $1/P_L$ selection criterion. As you can see from the figure, the algorithm correctly recognizes square objects as objects with maximum values of $1/P_L$. For elongated rectangular objects, these values are reduced.

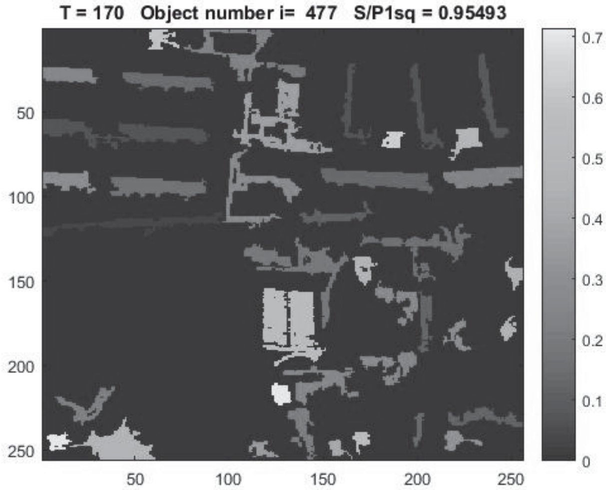
The results of processing for real images are shown in Fig. 9, where are the observations of the same area with the aircraft and satellite. The binarization thresholds for each object were set according to the maximum values of the criterion used, taking into account the preliminary removal of small objects. Despite the significant differences in the nature of the scenes and the statistics of observations, the algorithm quite successfully selects the same objects, which is especially important for solving the problems of comparison of heterogeneous images.

a)



b)

c)



d)

Fig. 9. Selection of objects on the test image by the elongation coefficient of the main axes. The inverse value was calculated

IV. Conclusion

The task of selecting objects of interest in an image is usually solved by segmenting the image, i.e. dividing it into non-intersecting areas. Selection of objects is carried out using various features that characterize the shape and texture of the object. As a rule, these stages are carried out sequentially, i.e. "areas of interest" are first defined, and only then the properties of these areas are analyzed. Traditional segmentation schemes do not directly use the properties or attributes of objects.

A method for setting the binarization threshold for each selected object after multi-threshold (multi-layer) processing based on the extremum of the geometric criterion is proposed. This criterion, in addition to the area of the object, can be the elongation coefficients of the perimeter or the main axis of the describing ellipse. These coefficients differ for round and extended objects and provide invariance of characteristics with changes in the image scale.

The methodology considered is rather simple with the algorithm having only a few free parameters that are easy to interpret. We note two important parameters that determine the boundary conditions, namely the minimum base size of the Sb object and the maximum P_S coefficient (or P_L coefficient), which will be taken into account in multilayer content analysis. A reasonable choice of both parameters allows you to speed up the algorithm, as well as eliminate high-frequency noise, represented by numerous small isolated objects many of them containing only one or more pixels. However, the proper selection of both parameters requires knowledge of the typical size of the object of interest, as well as the size of the image resolution, although in practice they are often adjusted during testing.

The results of selection of objects on typical noisy model and real television and infrared images have shown the efficiency and effectiveness of selection of extended spots and elongated objects of interest with minimal distortion of their boundaries at a fairly low signal/noise ratio. The performance of the proposed approach has been evaluated in comparison with previously reported binary integration method that had been implemented both numerically and analytically. Our results indicate that under comparable quantitative efficacy indicators the proposed approach clearly outperforms the binary integration method in terms of the object shape preservation.

We like to acknowledge the support of this work by the Ministry of Science and Higher Education of the Russian Federation via the basic state assignment of St. Petersburg Electrotechnical University, project No. 2.5475.2017/6.7.

References

[1] G. Gao, "Statistical modeling of SAR images: A Survey", *Sensors*, vol. 10, 2010, pp. 775-795.

[2] W. Zhou, A. Troy, "An Object-Oriented Approach for Analyzing and Characterizing Urban Landscape at the Parcel Level", *International Journal of Remote Sensing*, vol. 29, No. 11, 2008, pp. 3119-3135.

[3] V. Volkov, "Extraction of extended small-scale objects in digital images", *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci*., XL-5/ W6, 2015, pp. 87-93. Web: https://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XL-5-W6/87/2015/.

[4] M. Bogachev, V. Volkov, O. Markelov, E. Trizna, D. Baydamshina, V. Melnikov, R. Murtazina, P. Zelenikhin, L. Sharafutdinov, A. Kayumov, "Fast and simple tool for the quantification of biofilm-embedded cells sub-populations from fluorescent microscopic images", *PLoS One*. vol. 13. 2018, Iss. 3. P. e0192022. Web: http://journals.plos.org/plosone/article/file?id=10.1371/journal.pone.0193267&type=printable.

[5] M. Bogachev, V. Volkov, G. Kolaev, L. Chernova, I. Vishnyakov, A. Kayumov, "Selection and Quantification of Objects in Microscopic Images: from Multi-Criteria to Multi-Threshold Analysis", *Bionanoscience*, vol. 2018, pp.1-7. Web: https://doi.org/10.1007/s12668-018-0588-2.

[6] Volkov V.Y. Adaptive Extraction of Small Objects in Digital Images. Journal of the Russian Universities. Radioelectronics. 2017; (1):17-28. (In Russ.).

[7] H. Gu, Y. Han, Y. Yang, H. Li, Z. Liu, U. Soergel, T. Blaschke, S. Cui, "An efficient parallel multi-scale segmentation method for remote sensing imagery", *Remote Sensing,* 2018, 10(4), p. 590.

[8] J. Cheng, Y. Tsai, W. Hung, S. Wang, M. Yang, "Fast and accurate online video object segmentation via tracking parts", in *Proc. of the 2018 IEEE Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT USA, 2018, pp 7415-7424.

[9] M. Wang, "A multiresolution remotely sensed image segmentation method combining rainfalling watershed algorithm and fast region merging", *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci*., vol. XXXVII. Part B4. Beijing 2008.

[10] S. Arora, J. Acharya, A. Verma, P.K. Panigrahi, "Multilevel thresholding for image segmentation through a fast statistical recursive algorithm", *Pattern Recognition Letters,* vol. 29, 2008, pp. 119-125.

[11] J. Yang, Y. Yang, W. Yu, J. Feng, "Multi-threshold Image Segmentation based on K-means and Firefly Algorithm", *Proc. of 3rd Int. Conf. on Multimedia Technology.* vol. ICMT-13, 2013, Atlantis Press, pp. 134-142.

[12] P. Priyanka, K. Vasudevarao, Y. Sunitha, B.A. Sridhar, "Multi level fuzzy threshold image segmentation method for industrial applications", *IOSR Journal of Electronics and Communication Engineering* (IOSR-JECE), 2017, vol. 12, issue 2, ver. III (Mar.-Apr. 2017), pp. 06-17.

[13] O. Banimelhem, Y. Yahya, "Multi-thresholding image segmentation using genetic algorithm", *Proc. IPCV*, 2012.

[14] E. Cuevas, A. González, F. Fausto, D. Zaldívar, M. Pérez-Cisneros, "Multithreshold Segmentation by Using an Algorithm Based on the Behavior of Locust Swarms", *Hindawi Publishing Corporation. Mathematical Problems in Engineering,* vol. 2015, Article ID 805357.

[15] H. M. Finn, S. Johnson, "Adaptive Detection Mode with threshold control as a function of spatially sampled clutter level estimators", *RCA Rev*, vol. 29, No. 3, 1968, pp. 414-465.

[16] Dillard, G. "A moving-window detector for binary integration." IEEE Transactions on Information Theory 13.1 (1967): 2-6.