ICIC Express Letters ICIC International ©2020 ISSN 1881-803X
Volume **14**, Number **11**, November **2020** pp. 1041–1048

DETECTION OF BACKGROUND AREAS AND LOCALIZATION OF OBJECTS IN THE IMAGE

Badr Khalil Mahmoud El-Eissawi
and Victor Aleksandrovich Nesterenko*

Institute of Mathematics, Mechanics and Computer Sciences
South Federal University
Milchakova 8a., Rostov-on-Don 344090, Russia
*Corresponding author: neva09@mail.ru

Received January 2020; accepted April 2020

Abstract. *The paper describes the method of detecting background areas and localizing objects in the image. The background area is a large connected group of pixels with similar characteristics. To find the characteristics of the pixels, it is proposed to use Haar-like features. We use a modified Fuzzy Clustering Method (FCM) to search for groups of pixels that are close in the image and the feature space at the same time. We use the characteristics of the background areas to localize other objects in the image.*

Keywords: Image processing, Background area, Image segmentation, Localization of objects, Haar-like features, Fuzzy clustering

1. **Introduction.** Object localization problem is still a major challenge in pattern recognition [1]. If we know the characteristics of the object (the location of the eyes and lips on the face, the size and position of the signs on the license plate of the car, etc.), then we can localize the object by using the method of Viola-Jones [2] or the method of special points on the image [3]. However, when the characteristics of the object are unknown (searching for the objects of artificial origin on aerial photographs for example), these methods are usually not applicable. In this case, we can solve the localization problem in the following way: determine the characteristics of the background and consider the set of points differing from the background as objects in the image.

Separating an image into a background and objects placed on this background can be a challenge as objects can have different shapes, colors, brightness, and other characteristics. A method called background subtraction [4] is used to solve this problem. With this approach, the first step is to create a model of the image consisting only of the background and then identify the difference between the model and the real image to find objects. Often the background model comprises a static part of the image without moving objects [4] or a training image as a "blank scene" without objects [5].

The problem of splitting the image into various areas is almost similar to the image segmentation problem [6-9], but the only difference is that the background areas can contain foreign objects whose size is much smaller than the size of the background area.

In our work, we propose a two-component image model. We base the model on a hierarchy of typical scales of image parts. At the top-level of the hierarchy are large areas, which we will call the background areas of the image. The background region is large and homogeneous connected groups of pixels. Some images can contain more than one background area.

Objects will form up the next level of the scale hierarchy. We consider objects as a local part of the image that differs in characteristics from the background areas. The size of an object should be smaller than the size of the background area. The difference in the

DOI: 10.24507/icicel.14.11.1041

sizes of the object and the background area should be such that including or excluding an object from the background area that means it should have little effect on the background characteristics.

The lowest-level of the hierarchy contains elementary components of the image, and these are called textons or tokens. The size of these parts is usually a few pixels. The size of the objects must be larger than the size of the elementary tokens; therefore, the object is distinguishable among the tokens and cannot be lost in random fluctuations in the characteristics of the image pixels.

The proposed hierarchical model of the image is somehow conditional which means it might not work for some images. However, this model is intuitive and can be used for analyzing images, highlighting background areas and localizing objects.

This paper presents the method for analyzing and identifying the groups of pixels with similar characteristics. We assume that the image comprises background areas and objects. We apply a fuzzy clustering method to finding groups of connected pixels. Then we use the characteristics of the background areas to localize other objects in the image. Recently, more and more researchers are applying clustering methods to solving image processing problems. [10] contains a detailed description of the use of the K-means method for image segmentation. Authors of [10] use gray levels as a characteristic of image pixels. Authors of [11] use fuzzy clustering for segmentation of aerial photographs. In [11] authors use the levels of three spectral components as pixel characteristics. Other authors [6-9] also use the characteristics of a single-pixel or a small group of pixels. These characteristics describe the average properties of the image area without considering its structure. We propose to use Haar-like features as pixel characteristics. Haar-like functions give us the ability to use the collective characteristics of a group of pixels when analyzing an image. Collective pixel characteristics allow us to take or ignore individual image details and apply a hierarchical model of the image.

The presented method is the further development of the methods proposed in [12,13]. We described the image as a single background area and the objects on it. We used the averaged characteristics of the image as the characteristics of the background area. In this paper, we use several background areas with different characteristics. We chose a different set of pixels characteristics compared to [12,13]. These changes improve on how the proposed method performs.

This paper is organized as follows. In Section 2 we describe the method for determining the background areas and discuss the characteristics of the pixels which are required for this method. In Section 3 we give an approach to finding the number of background areas. In Section 4 we offer examples of detection of background areas in images. Here we use test images and real aerial photography. In Section 5 we use the results of Section 4 to localize objects in the background areas. The conclusion is given in Section 6.

2. **Detection of the Background Area in the Image.** Let us define the background part of the image as large, connected areas with similar pixel characteristics. The test image 800×400 (Figure 1) contains two background areas: the left and the right parts of the image. We use textures from the photographic album of textures by Brodatz [14] to form images of Figure 1 (D17, D77). The size of the background areas of the image is 400×400 pixels. Also, the image contains two objects that are not the background. We assume that the size of the background area is much larger than the size of the object. In the above case, the dimensions of the objects are 88×88 pixels. Adding or removing an object from the background area should have little effect on the background characteristics. The sizes of tokens (textons) are 7 and 11 pixels for the left and right parts of the image.

We will use Haar-like functions [15] as characteristics that determine whether the given pixel belongs to a background or an object. Let $B(i, j)$ denote the brightness of the pixel

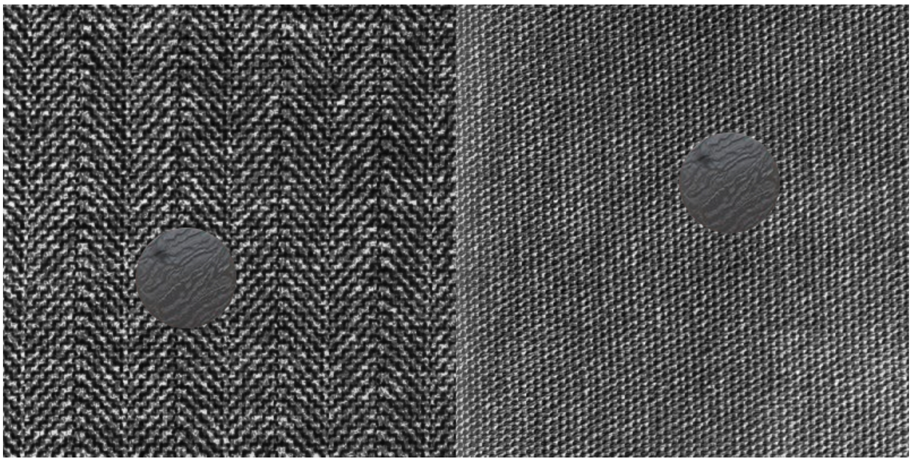

FIGURE 1. Test image with background areas and objects

with the coordinates (i, j). The pixel (i_0, j_0) will have a set of characteristics calculated in the neighborhood $\Omega(i_0, j_0)$ using the masks $s_m(i, j)$:

$$P_m(i_0, j_0) = \sum_{(i,j)\in\Omega(i_0,j_0)} s_m(i - i_0, j - j_0)B(i, j) \tag{1}$$

The area $\Omega(i_0, j_0)$ is rectangle $S_x \times S_y$ with the center (i_0, j_0). Figure 2 shows the set of masks.

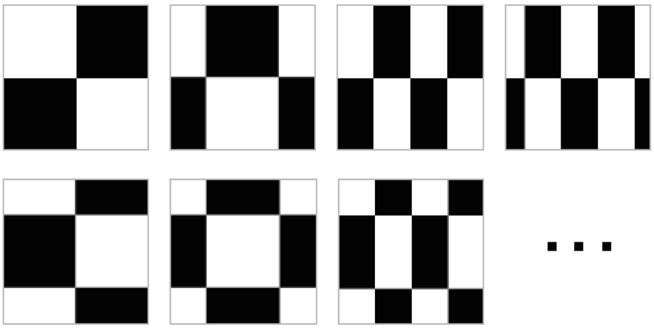

FIGURE 2. Samples of Haar masks

The summation in the expression (1) is equal to the evaluation of the integral

$$h_{kl}^{ab} = \int_{x-S_x/2}^{x+S_x/2} \int_{y-S_y/2}^{y+S_y/2} dsdtB(s,t)f_k^{(a)}(s-x)f_l^{(b)}(t-y) \tag{2}$$

where instead of the mask $s_m(i, j)$ we use the functions $f_k^{(a)}(s)f_l^{(b)}(t)$. The functions $f_n^{(a)}(x)$ has the form shown in Figure 3.

If in Formula (2) we replace the meander functions with $\sin(x)$ and $\cos(x)$, we get the formula of the Fourier transform. Applying the Fourier transform instead of (2) will yield results similar to the results of this paper. Using meander functions corresponds to the use of Haar masks in Formula (1). This allows us to use the SAT (Summed-Area Tables [16]) method and simplifies calculations.

The value

$$w_{kl} = \left(h_{kl}^{11}\right)^2 + \left(h_{kl}^{21}\right)^2 + \left(h_{kl}^{12}\right)^2 + \left(h_{kl}^{22}\right)^2 \tag{3}$$

is similar to the density of the energy spectrum for the Fourier transform. In this paper we will use the values H_m as the set of pixel characteristics.

$$H_m = \sqrt{klw_{kl}} \qquad \{m\} = \{k\} \otimes \{l\} \tag{4}$$

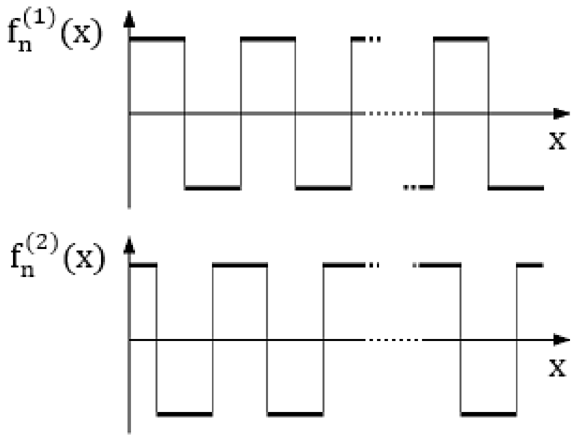

FIGURE 3. The meander functions

When we remove the characteristics for odd $(k \cdot l)$ from the set of H_m, it makes H_m independent from the brightness of an image. The characteristics H_m identify the area of the image according to the typical sizes of the tokens in that area.

Thus, we transform a two-dimensional image surface into a space H of characteristics of dimension M. The distance

$$d\left(H^{(1)}, H^{(2)}\right) = \sqrt{\sum_{m=1}^{M} \left(H_m^{(1)} - H_m^{(2)}\right)^2} \tag{5}$$

permits us to estimate the similarity of the characteristics of two pixels. When the distance $d\left(H^{(1)}, H^{(2)}\right)$ inside the pixel group is small and the number of points in the group is large, then we identify this set as the background. We will solve the problem of finding these groups of pixels by cluster analysis methods. The test image (Figure 1) shows that the pixels must form two clusters corresponding to the left and the right sides of the image.

We use the above-mentioned Formulas (1)-(4) to calculate the characteristics of the pixels of the test image Figure 1. As the clusters from different background areas may overlap [12], we use Fuzzy Classifier Means (FCM) method [17] to find clusters corresponding to different background areas in the image.

In the FCM algorithm, the matrix μ_{ki} specifies the probability that an element $H^{(i)}$ belongs to a cluster $C^{(k)}$:

$$\mu_{ki} = \frac{A_i}{d\left(C^{(k)}, H^{(i)}\right)^{1/q}}$$

where q determines the level of cluster fuzziness, A_i – the normalizing factor: $A_i = \sum_k d\left(C^{(k)}, H^{(i)}\right)^{1/q}$.

Background pixels in an image must have neighbors with similar properties. We change the probability μ_{ki} to consider the dependence on neighboring pixels:

$$\mu_{ki} = \frac{\tilde{A}_i \left(1 + \alpha N_{ki}\right)}{d\left(C^{(k)}, H^{(i)}\right)^{1/q}} \tag{6}$$

where N_{ki} is the number of pixels from the cluster $C^{(k)}$ nearby to the element $H^{(i)}$, $3 \leq N_{ki} \leq 8$. The α parameter determines the influence of neighbors. The values of the parameters q and α depend on the individual task and we should take their values to get a good clustering result.

3. **Determining the Number of Background Areas.** The number of background areas detected is equal to the number of clusters in the characteristic space. The number of clusters can be determined by the clustering quality criterion. We use the value of the square deviation D_k in the cluster $C^{(k)}$ for this purpose [18]:

$$D_k = \sum_{\left\{i: H^{(i)} \in C^{(k)}\right\}} d^2\left(H^{(i)} - M^{(k)}\right)$$

where $M^{(k)}$ is the geometric center of the cluster $C^{(k)}$. We introduce the criterion of proximity of the clusters:

$$K = \frac{D_{ab} - (D_a + D_b)}{(D_a + D_b)} \tag{7}$$

Here D_{ab} is the quadratic deviation in the merged cluster $C^{(a,b)} = C^{(a)} \cup C^{(b)}$.

Let us set the number of clusters more than necessary. Let $C^{(a)}$ and $C^{(b)}$ be the closest clusters in the space of characteristics. If the criteria value K is small, then these clusters are part of a single shared cluster. In this case, we can cut the number of clusters and repeat the clustering procedure.

Table 1 contains values for criterion K (7) for a different number of clusters in the test image in Figure 1.

Table 1. Variation of criterion K depending on the number of clusters

Number of clusters	Criterion K
5	0.003
4	0.009
3	0.055
2	0.156

The above data shows that the number of clusters and background areas in this case should be equal to 2.

4. **Examples of Detection of the Background Areas.** The result in Figure 4 shows that the proposed method highlights two background areas in the test image.

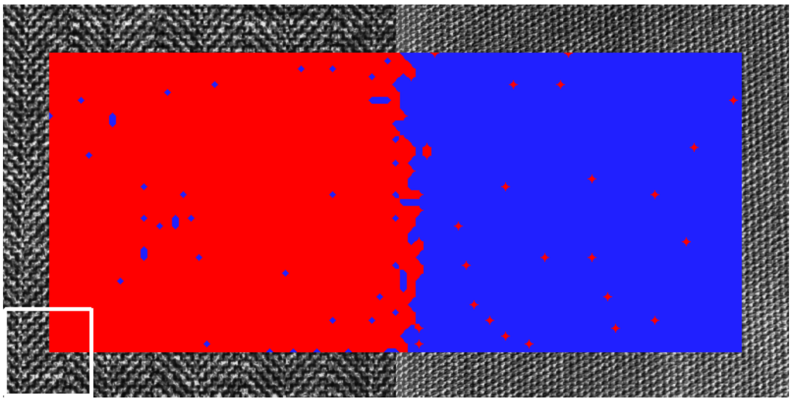

Figure 4. (color online) The result of selection background areas in the test image

We used the following parameter values for this image sample:

- The sizes of the rectangular area $\Omega(i, j)$ for the calculation (1), (2) are $S_x = 90$ and $S_y = 90$.
- The set values of the half-periods for the functions of the meander $f_n^{(a)}(x)$ are $\{1, 2, 3, 5, 6, 9, 10, 15, 18, 30, 45\}$.

- Dimension of the characteristic space H is $M = 85$ (we exclude the combinations $\{1,3,5,9,15,45\} \otimes \{1,3,5,9,15,45\}$).
- Parameter values in (6) are $q = 2.0$ and $\alpha = 0.35$.

Using the method Summed-Area Tables [16] and a reasonable choice of sparsity of points for the calculation of characteristics allows performing efficient calculations, despite the large number of characteristics $M = 85$. A quantity of characteristics depends on the size 90×90 region $\Omega(i,j)$.

The sizes of the objects in Figure 1 specify the size of the region $\Omega(i,j)$. The choice of size of a region $\Omega(i,j)$ and the value of the parameter α in Equation (6) allows excluding foreign objects when identifying the background areas. The white square in the lower left corner of Figure 4 shows the $\Omega(i,j)$ area at 90×90 pixels.

The parameters characterizing the proportion of correctly and incorrectly classified pixels are: Sensitivity (true positive rate) = 0.97, Miss Rate (false negative rate) = 0.03. These values can serve as an estimate of the accuracy of the method. However, these values may vary depending on similarity of the characteristics of different background areas, the presence of objects and because of the probabilistic nature of the method used fuzzy clustering FCM.

The set of characteristics (4) used to find the background allows us to distinguish areas according to the characteristic dimensions of the image details in the background areas and the distribution of their size by value. This choice of characteristics allows to find many background areas like those shown in Figure 4. Experimental testing for other images confirms this.

Figure 5 shows the results of applying the proposed method to aerial photograph [19].

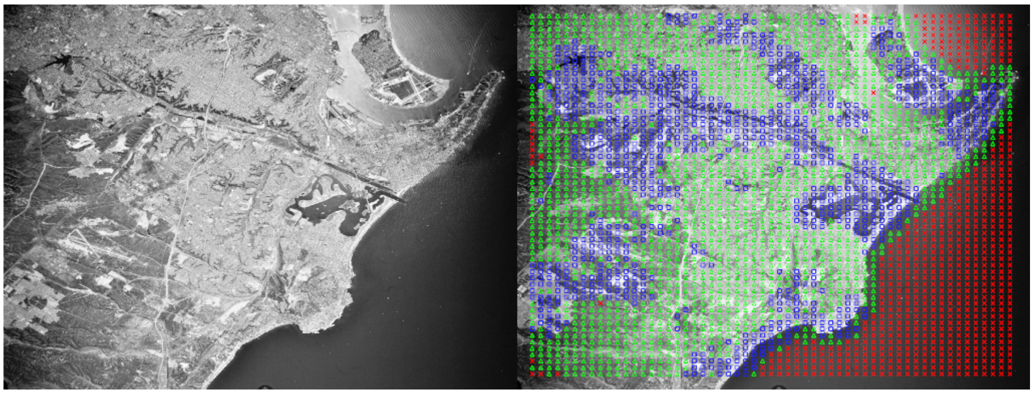

FIGURE 5. (color online) The result of selection background areas in the aerial photograph

5. **Localizing Objects in the Image.** After analyzing the background areas in the image, we can solve the problem of localizing objects in these areas of the background. To do this, we use the pixel characteristics (4) within the found background area.

According to the proposed hierarchical model the size $\Omega(i,j)$ in Formulas (1) and (2) should be smaller than the sizes of the objects and larger than the scale of the tokens. This size is 24×24 pixels for the test image Figure 1 (white square in the lower left corner in Figure 6). The set values of the half-periods for the functions of the meander $f_n^{(a)}(x)$ are 1, 2, 3, 4, 6, 8, 12 and the dimension of the characteristic space is $M = 45$ (we exclude the odd combinations $\{1,3\} \otimes \{1,3\}$). To separate objects and background areas of the image, we again use Haar-like features and the fuzzy clustering method as described earlier in Section 2.

We show the result of separating the object and background in Figure 6.

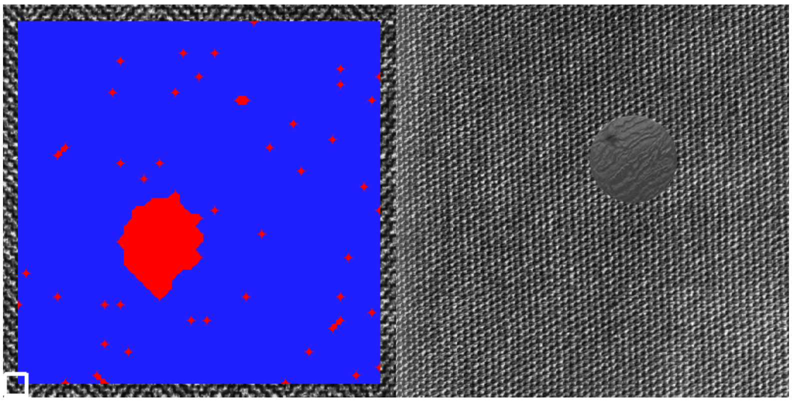

Figure 6. (color online) The result of localizing the object in the background area

We can use another way to highlight objects on the found background. We will highlight those pixels for which the characteristics of H_m (4) differ greatly from the mean values of the characteristics of the background area. To calculate the difference between pixels and the background we use a metric:

$$d\left(H,\ H^{(bkg)}\right) = \sqrt{\sum_{m=1}^{M}\left(\frac{H_m - H_m^{(bkg)}}{H_m + H_m^{(bkg)}}\right)^2}$$

where H is the Haar feature set of the pixel, $H^{(bkg)}$ is the Haar feature set of the background area. If the distance d is greater than the specified threshold value:

$$d\left(H,\ H^{(bkg)}\right) > T_{obj}$$

then we believe this pixel does not belong to the background area. If the threshold criterion T_{obj} is small, we will see some background pixels as object pixels. Vice versa, if the criterion T_{obj} is high, we will observe some pixels of the object as background pixels. The threshold value T_{obj} depends on the image being studied and may vary for various tasks.

Figure 7 shows the result of this localization of the object. The red color marks pixels that are disparate from the background. For the result in Figure 7 we used the value $T_{obj} = 2.75$.

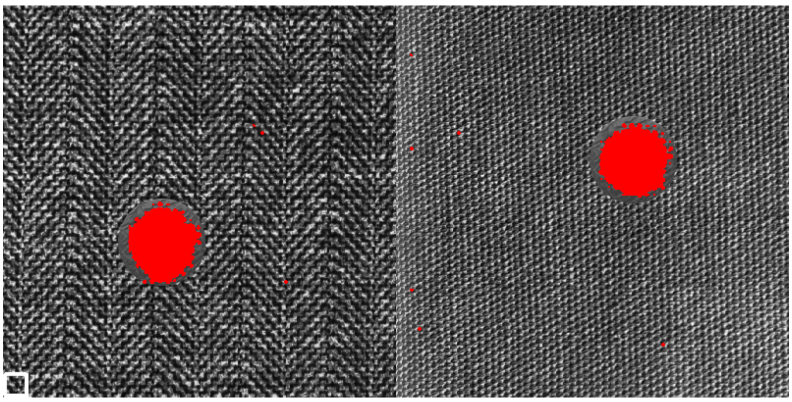

Figure 7. (color online) The result of localizing objects using a different method

Two methods of object localization have been proposed in this paper. While identifying an object and establishing its attributes is beyond this paper and other methods should solve it.

6. **Conclusion.** The paper proposes the method for detection of background areas and localization of objects in the image. The method is based on detection of closely spaced, connected groups of pixels with similar characteristics. The results of the application of the proposed method to the test image and the photograph are given.

The limitations of this method include the following:

- the image must consist of background areas and foreign (relative to the background) objects;
- the characteristics should not change essentially much in the background areas;

The obtained results say the possibility of using the proposed method to identify background areas in the image. We can use these results to further localize objects to the selected background.

REFERENCES

[1] A. Andreopoulos and J. K. Tsotsos, 50 years of object recognition: Directions forward, *Computer Vision and Image Understanding*, vol.117, pp.827-891, 2013.

[2] P. A. Viola and M. J. Jones, Rapid object detection using a boosted cascade of simple features, *Proc. of 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp.511-518, 2001.

[3] C. Harris and M. Stephens, A combined corner and edge detector, *Proc. of the 4th Alvey Vision Conference*, pp.147-151, 1988.

[4] M. Piccardi, Background subtraction techniques: A review, *IEEE International Conference on Systems, Man and Cybernetics*, pp.3099-3104, 2004.

[5] B. Han and L. S. Davis, Density-based multifeature background subtraction with support vector machine, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.34, no.5, pp.1017-1023, 2012.

[6] N. R. Pal and S. K. Pal, A review on image segmentation techniques, *Pattern Recognition*, vol.26, pp.1277-1294, 1993.

[7] D. Liu, B. Soran, G. Petrie and L. Shapiro, A review of computer vision segmentation algorithms, *Lecture Notes*, 2012, https://courses.cs.washington.edu/courses/cse576/12sp/notes/remote.pdf, Accessed on October 9, 2019.

[8] N. M. Zaitouna and M. J. Aqelb, Survey on image segmentation techniques, *Procedia Comput. Sci.*, vol.65, pp.797-806, 2015.

[9] Y. Song and H. Yan, Image segmentation techniques overview, *2017 Asia Modelling Symposium (AMS)*, 2017.

[10] Z. Khan, J. Ni, X. Fan and P. Shi, An improved K-means clustering algorithm based on an adaptive initial parameter estimation procedure for image segmentation, *International Journal of Innovative Computing, Information and Control*, vol.13, no.5, pp.1509-1525, 2017.

[11] M. M. Trivedi and J. C. Bezdek, Low-level segmentation of aerial images with fuzzy clustering, *IEEE Trans. Systems, Man and Cybernetics*, vol.16, no.4, pp.589-598, 1986.

[12] V. A. Nesterenko, Localization of foreign objects in the image, *Proc. of the XXIV Conference on Modern Information Technologies: Trends and Prospects of Development*, Southern Federal University, Rostov-on-Don, pp.145-148, http://inftech.uginfo.sfedu.ru/sites/default/files/Abstr%20SITO%202017.pdf, Accessed on August 26, 2020 (in Russian).

[13] V. A. Nesterenko, The search for background areas in an image, *Electronic Scientic Journal Information Processes*, vol.16, no.3, pp.291-300, 2016 (in Russian).

[14] P. Brodatz, *Textures: A Photographic Album for Artists and Designers*, Dover Publications, New York, 1966, http://www.ux.uis.no/tranden/brodatz.html, Accessed on July 25, 2018.

[15] C. P. Papageorgiou, M. Oren and T. Poggio, A general framework for object detection, *Proc. of the 6th International Conference on Computer Vision*, pp.555-562, 1998.

[16] F. C. Crow, Summed-area tables for texture mapping, *SIGGRAPH '84: Proc. of the 11th Annual Conference on Computer Graphics and Interactive Techniques*, pp.207-212, 1984.

[17] J. C. Bezdek, R. Ehrlich and W. Full, FCM: The fuzzy c-means clustering algorithm, *Computers & Geosciences*, vol.10, nos.2-3, pp.191-203, 1983.

[18] R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*, Wiley-Interscience, London, 1973.

[19] *The USC-SIPI Image Database*, http://sipi.usc.edu/database, Accessed on July 25, 2018.