

Correlation properties of sea surface images on video stream frames

Daria A. Chernomorets^{a,*}, Victor Golikov^b, Tatyana N. Balabanova^a, Ekaterina I. Prokhorenko^a, Evgeniya V. Bolgova^a, Andrey A. Chernomorets^a

^aBelgorod State University 85, Pobedy St., Belgorod, 308015, Russia

^bUniversidad Autónoma del Carmen, Calle 56 No. 4 Esq Avenida Concordia Col, Benito Juárez, Del Carmen, Campeche, 24180, Mexico

(Communicated by Seyed Hossein Siadati)

Abstract

The work is devoted to the characteristic properties analysis of the sea surface image on video stream frames, which makes it possible to make a decision about the detection of object images on them in the absence of a priori information about objects, which is important in the development of navigation systems, security and surveillance systems. In this paper, it is proposed to analyze the values of the normalized cross-correlation coefficients of a given image fragment with higher dimension fragments on sequential video stream frames to identify the distinctive properties of the agitated sea surface and floating objects images. The computational experiments were carried out and have shown that the analysis of the results of calculating the frame fragments cross-correlations allows us to estimate the amount of displacement and distortion of a given image fragment. The results of computational experiments demonstrate the presence of differences in the values of the corresponding cross-correlation coefficients for the sea surface images with different agitation degrees, containing and not containing an image of the object. Based on the analysis of the proposed correlation properties of the sea surface images, a decisive rule for selecting fragments of frames containing an image of an object is formulated, the use of which, in many cases, allows to detection of fragments of object images correctly.

Keywords: image, agitated sea surface, video stream frames, normalized cross-correlation, image fragment distortion, image fragment offset, object detection

2020 MSC: 54H30

1 Introduction

Solving the problem of identifying images of floating objects on the video stream frames containing information about the sea surface is relevant in the development of navigation systems, security and surveillance systems, during search and rescue operations, etc. Today a significant research amount of various sea surface and object images properties on a frame sequence of in order to computerize the process of observing the sea surface is being carried out [4, 18].

*Corresponding author

Email addresses: chernomorets_d@bsuedu.ru (Daria A. Chernomorets), vgolikov@pampano.unacar.mx (Victor Golikov), tatyanan.balabanova@gmail.com (Tatyana N. Balabanova), ekaterinai.prokhorenko@gmail.com (Ekaterina I. Prokhorenko), evgeniyav.bolgova@gmail.com (Evgeniya V. Bolgova), andriya.chernomorets@gmail.com (Andrey A. Chernomorets)

When solving the problem of detecting objects on the sea surface, various methods analyze various characteristic properties (features) of images of the sea surface and objects floating on it, in order to identify differences sufficient to make a decision about the location of the object.

In order to obtain information about the location of object images in the frames when analyzing video recordings, various methods are used in which the following features are analyzed: object features formed on the basis of the their contours selection obtained as a result of calculating the first and second image derivatives [6]; features based on the calculation of statistical image characteristics, such as, histogram, arithmetic mean, variance, median, etc. [5, 10, 13]; features based on the description of the object structure and its key points (descriptors), such as angles, edges, etc. [1, 2]; fractal features, obtainment of which is based on the principle of objects self-similarity [14]; features that allow to evaluate the movement of an object, obtained, for example, as a result of background selection based on Gaussian mixtures, constructing a velocity vector when calculating the optical flow, etc. [9, 12]; features obtained on the basis of linear transformations, such as the Fourier transformation, cosine transformation, Walsh-Hadamard and Haar transformations, etc., used to identify the spectral characteristics of images both in the spatial domain of a single frame and in the time domain when analyzing elements of sequential frames, etc. [3, 11, 15, 16, 17].

However, most of the known methods use a priori information about the properties of the sea surface and objects images. In this paper, authors analyze the characteristic properties of the agitated sea surface and objects images on the video stream frames, which allow to make a decision about the detection of objects images on them in the absence of a priori information about the objects and the sea surface agitation degree.

2 Method

Let consider the properties of the agitated sea surface and objects images on sequential video stream frames, based on the analysis of the values of the normalized cross-correlation coefficients of video stream frames fragments.

Consider the image of the object on the adjacent video stream frames. The image of an object in a frame sequence is usually characterized by no significant distortion, as well as no significant shift and rotation relative to the position of the object in the previous frame.

To estimate the amount of distortion and offset of the object image, authors propose to calculate the normalized cross-correlation of object image fragment on the selected frame with a given area on the following frames. As a result, the maximum value of the cross-correlation coefficients, in most cases, indicates the area of the next frame corresponding to the new position of the analyzed fragment of the object image.

The cross-correlation coefficient maximum value indicates the degree of difference-similarity of the found area in the adjacent frame and the original object image fragment. With a sufficiently large value of this coefficient, for example, not less than 0.9, it can be assume that the area found in the next frame does not significantly differ from the original fragment of the object image and corresponds to the position of the analyzed fragment in the next frame. The position of the area in the next frame corresponding to this coefficient can be used to estimate the amount of the original image fragment offset.

The normalized cross-correlation coefficient of the matrix T containing the pixel values of the analyzed object image fragment, and the matrix Φ corresponding to the pixels of the frame following the frame containing the analyzed object image fragment, under the condition that all elements of the matrix T correspond to the specified elements of the matrix F , is calculated based on the following relation [7, 8]:

$$r_{uv} = \frac{\sum_{i=1}^{M_1} \sum_{j=1}^{M_2} (f_{i+v-1, j+u-1} - a_{uv})(t_{ij} - a_T)}{\sqrt{\sum_{i=1}^{M_1} \sum_{j=1}^{M_2} (f_{i+v-1, j+u-1} - a_{uv})^2 \sum_{i=1}^{M_1} \sum_{j=1}^{M_2} (t_{ij} - a_T)^2}}, \quad (2.1)$$

$$u = 1, 2, \dots, N_1 - M_1 + 1,$$

$$v = 1, 2, \dots, N_2 - M_2 + 1,$$

where $T = (t_{ij})$ – image (matrix) T , containing the analyzed object image fragment, size of $M_1 \times M_2$ pixels, a_T mean of matrix elements T ,

$$a_T = \frac{1}{M_1 M_2} \sum_{i=1}^{M_1} \sum_{j=1}^{M_2} t_{ij}, \quad (2.2)$$

$\Phi = (f_{ij})$ – image (matrix) Φ of the frame following the frame containing the analyzed fragment of the object image, size of $N_1 \times N_2$ pixels,

$$N_1 \geq M_1, \quad N_2 \geq M_2,$$

a_{uv} – mean of the submatrix of the matrix F elements corresponding to the given position of the matrix T,

$$a_{uv} = \frac{1}{M_1 M_2} \sum_{i=1}^{M_1} \sum_{j=1}^{M_2} f_{i+v-1, j+u-1}. \tag{2.3}$$

The values of the normalized cross-correlation coefficients vary in the interval [-1,1]. The coordinates of the image fragment are the values of the indices image pixel (the numbers of the corresponding row and column) located in the upper left corner of the rectangular image area of the frame that contains the analyzed image fragment.

Consider a sequence of images (frame fragments) $\Phi_1, \Phi_2, \dots, \Phi_K$, size of $N_1 \times N_2$ pixels, that have the same coordinates on the video stream frames.

In the image Φ_1 select the image (area) T, size of $M_1 \times M_2$ pixels, the upper left pixel of which corresponds to the image Φ_1 with coordinates $(u^{(1)}, v^{(1)})$. It is obvious that the maximum value of the cross-correlation coefficient (2.1) of the images T and Φ_1 is equal to 1. Define this value $r^{(1)}$, $r^{(1)} = 1$.

Calculate the values of the cross-correlation coefficient (2.1) of the images T and Φ_2 . Define $r^{(2)}$ – the maximum value of the cross-correlation coefficient (2.1) of the images T and Φ_2 , $(u^{(1)}, v^{(1)})$ – the indices corresponding to this value in expression (2.1).

Calculate on the basis of expression (2.1) values $r^{(k)}$, $(u^{(k)}, v^{(k)})$, $k = 3, \dots, K$, or pairs of images T and Φ_3, \dots, Φ_K .

Values $(u^{(k)}, v^{(k)})$, $k = 2, 3, \dots, K$, determine the position of the areas $T^{(k)}$, $k = 2, 3, \dots, K$, size of $M_1 \times M_2$ pixels, in the images $\Phi_2, \Phi_3, \dots, \Phi_K$, which correspond to the image T taking into account the value of the cross-correlation coefficient (Figure 1).

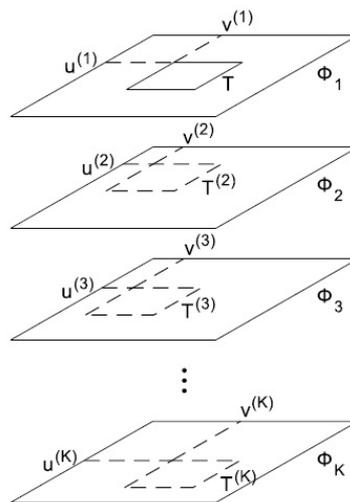


Figure 1: Areas on the sequence of frame fragments $\Phi_1, \Phi_2, \dots, \Phi_K$, corresponding to the image T

The analysis of the values of the sequences $\{r^{(k)}\}$, $\{(u^{(k)}, v^{(k)})\}$, $k = 1, 2, \dots, K$, allows to study the properties of sea surface images on the video stream frames.

3 Results

Consider computational experiments that demonstrate differences in the sequences of maximum values of the cross-correlation coefficients $\{r^{(k)}\}$ $k = 1, 2, \dots, K$, for the sea surface images with and without an object image.

To carry out computational experiments, two videos were selected from open Internet sources, corresponding to various degrees of the sea surface agitation. The first frames of the selected videos; Video1 and Video2 are shown in Figure 2.

When carrying out calculations, the value of the parameter K, the number of analyzed frame fragments, was chosen equal to 11, $K=10$.

The analyzed area T is chosen as a square shape, size of 10×10 pixels, $M_1 = M_2 = 10$.

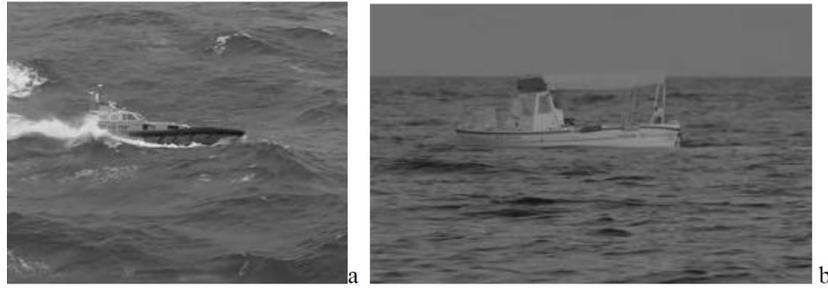


Figure 2: Fragments of analyzed videos: a – Video1, b – Video2

The sizes of frame fragments $\Phi_1, \Phi_2, \dots, \Phi_K$ or which it is proposed to calculate the cross-correlation with the area T , are chosen equal to 28×28 pixels, $N_1 = N_2 = 28$.

Consider the results of calculating the cross-correlation of video Video1 frame fragments (Figure 2a) containing the image of the object.

Figure 3 shows the fragments of frames $\Phi_1, \Phi_5, \Phi_{10}$ (highlighted square areas) of the video Video1, in which the image of the boat is located. Figure 3a also shows the analyzed area T , located in the center of the area Φ_1 .

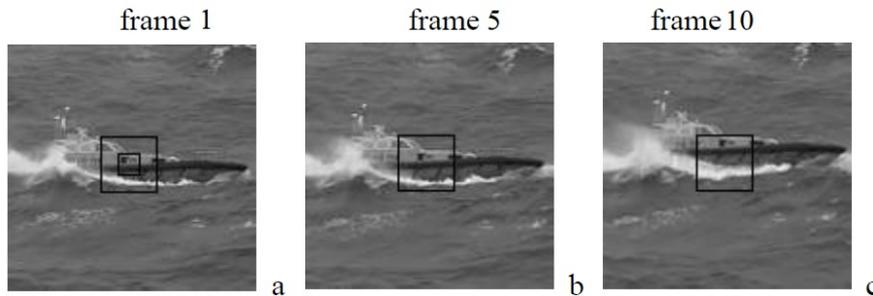


Figure 3: Fragments of frames $\Phi_1, \Phi_5, \Phi_{10}$ of video Video1 with object

In the images shown in Figure 3, you can see a slight offset of the boat image in the areas Φ_5, Φ_{10} relative to the area Φ_1 .

Table 1 shows the maximum values of the cross-correlation coefficient of the area T and the areas $\Phi_2, \Phi_3, \dots, \Phi_{10}$, corresponding to video Video1 frame fragments containing the object.

Table 1: The maximum cross-correlation coefficient of video Video1 frame fragments containing the object

Frame fragment number, k	Row number, $u^{(k)}$	Column number, $v^{(k)}$	Coefficient, $r^{(k)}$
1	10	10	1.000
2	10	10	0.992
3	10	10	0.991
4	9	10	0.902
5	9	10	0.981
6	8	10	0.977
7	6	10	0.917
8	5	11	0.969
9	3	11	0.909
10	2	12	0.928

The data in Table 1 show that the coordinates of the object corresponding to the area T in the first frame change slightly on different frames, that is, the object in the video moves. In this case, the image of the object changes slightly, which is indicated by a slight change in the cross-correlation coefficient.

Consider the results of calculating the cross-correlation of video frame fragments Video1 that do not contain the image of the object.

Figure 4 shows the fragments of frames $\Phi_1, \Phi_5, \Phi_{10}$ (highlighted square areas) of the video Video1, in which the image of the boat is not located. Figure 4a also shows the analyzed area T , located in the center of the area Φ_1 .

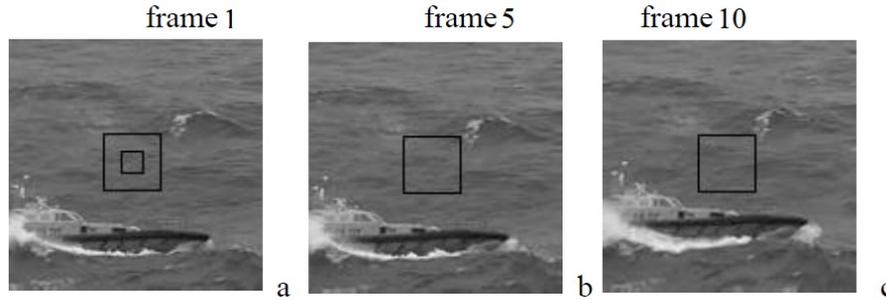


Figure 4: Fragments of frames $\Phi_1, \Phi_5, \Phi_{10}$ of video Video1 without object

Table 2 shows the maximum values of the cross-correlation coefficient of the area T and the areas $\Phi_2, \Phi_3, \dots, \Phi_{10}$, corresponding to video, Video1 frame fragments that do not contain the object.

Table 2: The maximum cross-correlation coefficient of video Video1 frame fragments that do not contain the object

Frame fragment number, k	Row number, $u^{(k)}$	Column number, $v^{(k)}$	Coefficient, $r^{(k)}$
1	10	10	1.000
2	10	9	0.864
3	11	8	0.926
4	11	7	0.892
5	11	6	0.810
6	11	5	0.711
7	10	4	0.718
8	9	4	0.700
9	8	4	0.700
10	7	3	0.697

The data in Table 2 show that the image of the sea surface area changes significantly, as indicated by a significant change in the cross-correlation coefficient when analyzing images on different frames. In this case, the coordinates of the corresponding detected area change slightly.

Consider the results of calculating the cross-correlation of video frame fragments Video2 (Figure 2b) containing the image of the object.

Figure 5 shows the fragments of frames $\Phi_1, \Phi_5, \Phi_{10}$ (highlighted square areas) of the video Video2, in which the image of the boat is located. Figure 5a also shows the analyzed area T, located in the center of the area Φ_1 .

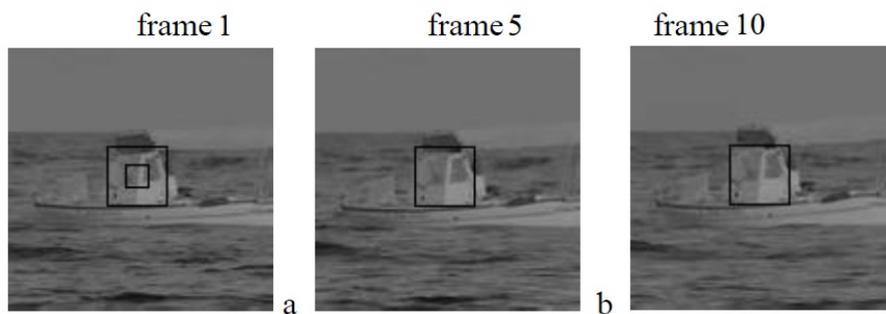


Figure 5: Fragments of frames $\Phi_1, \Phi_5, \Phi_{10}$ of video Video2 with object

In the images shown in Figure 5, there is practically no offset of the boat image in the areas Φ_5, Φ_{10} relative to the area Φ_1 . Table 3 shows the maximum values of the cross-correlation coefficient of the area T and the areas $\Phi_2, \Phi_3, \dots, \Phi_{10}$, corresponding to video Video2 frame fragments containing the object.

The data in Table 3 show that the coordinates of the object corresponding to the area T in the first frame change slightly on different frames, that is, the object in the video almost does not move. In this case, the image of the object changes slightly, which is indicated by a slight change in the cross-correlation coefficient. Consider the results of calculating the cross-correlation of video Video2 frame fragments that do not contain the image of the object.

Table 3: The maximum cross-correlation coefficient of video Video2 frame fragments containing the object

Frame fragment number, k	Row number, $u^{(k)}$	Column number, $v^{(k)}$	Coefficient, $r^{(k)}$
1	10	10	1.000
2	10	10	0.993
3	10	10	0.993
4	10	10	0.985
5	10	10	0.958
6	10	10	0.947
7	10	10	0.947
8	9	10	0.956
9	9	10	0.958
10	9	10	0.921

Figure 6 shows the fragments of frames $\Phi_1, \Phi_5, \Phi_{10}$ (highlighted square areas) of the video Video2, in which the image of the boat is not located. Figure 6a also shows the analyzed area T, located in the center of the area Φ_1 .

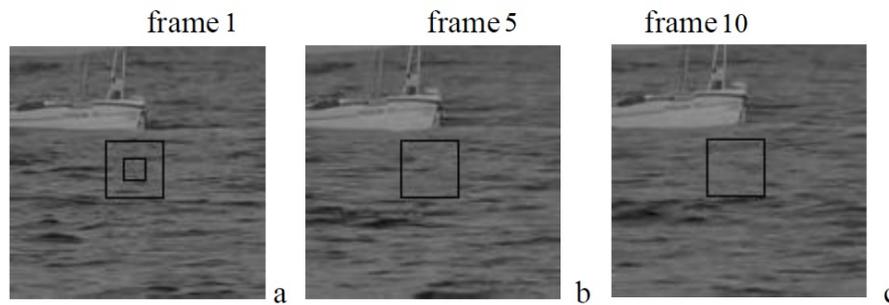
Figure 6: Fragments of frames $\Phi_1, \Phi_5, \Phi_{10}$ of video Video2 without object

Table 4 shows the maximum values of the cross-correlation coefficient of the area T and the areas $\Phi_2, \Phi_3, \dots, \Phi_{10}$, corresponding to video Video2 frame fragments that do not contain the object.

Table 4: The maximum cross-correlation coefficient of video Video2 frame fragments that do not contain the object

Frame fragment number, k	Row number, $u^{(k)}$	Column number, $v^{(k)}$	Coefficient, $r^{(k)}$
1	10	10	1.000
2	10	10	0.824
3	10	10	0.561
4	10	10	0.447
5	9	9	0.474
6	14	19	0.442
7	7	7	0.559
8	7	9	0.550
9	7	9	0.572
10	7	10	0.454

The data in Table 4 show a significant change in the cross-correlation coefficient when analyzing images on different frames, which corresponds to significant changes in the sea surface images. In this case, the coordinates of the corresponding detected area change abruptly.

4 Discussion

The data given in Tables 1–4 show that the analysis of the cross-correlation coefficient values of the image areas on the frame sequence, as well as the analysis of changes in the coordinate values corresponding to the maximum value of the cross-correlation coefficient, allows to identify differences in the correlation properties of the sea surface images that contain and do not contain object images.

Figure 7 shows the plots of the maximum values of the cross-correlation coefficients given in Tables 1–4 for the image fragments on the frames of videos Video1 and Video2, on which the boat image is present or absent.

The plots given in Figure 7 clearly show that the normalized cross-correlation coefficient retains values greater than 0.9 when analyzing frame fragments containing the boat image, which is associated with minor changes in its image

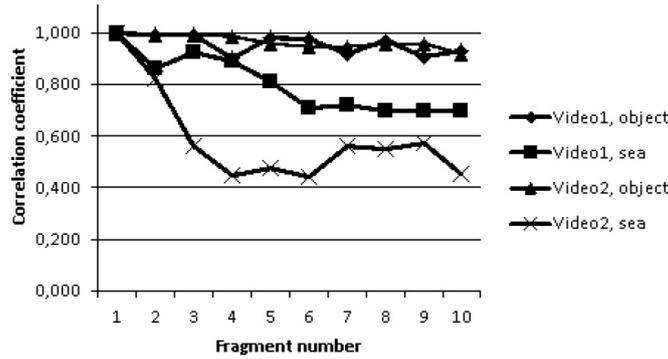


Figure 7: Cross-correlation coefficients of frame fragments

on adjacent video frames. The plots also show a significant decrease in the values of the cross-correlation coefficient for fragments of sea surface frames that do not contain a boat image, which is caused by significant changes in the images on the frame sequence (change in time).

Consequently, the results of calculating the normalized cross-correlation of the corresponding video frame fragments determine the correlation properties of the sea surface images, which can be used to develop methods for detecting floating objects.

Based on the analysis of the corresponding coefficients values, authors formulate the following decisive rule (preliminary, requiring clarification) for selecting frame fragments containing an object image, for example,

$$\ll \text{If } r^{(k)} > 0.9, \text{ then the frame fragment contains an object image} \gg . \tag{4.1}$$

The result of applying the decision rule (4.1) when analyzing images on video frames of Video1 and Video2 is shown in Figure 8. In Figure 8, the squares mark the areas of the frames in which, according to the decision rule (4.1), the object image is assumed to be present.

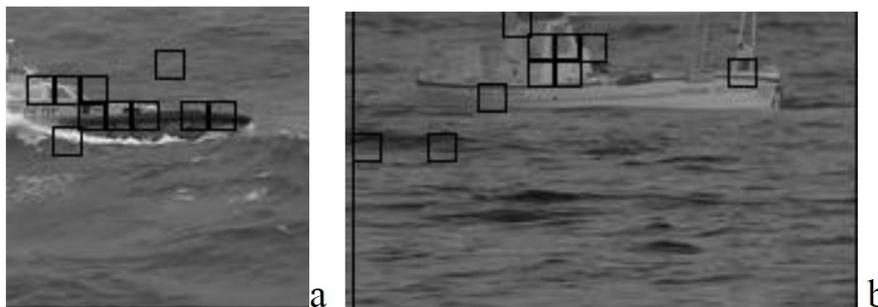


Figure 8: Applying the decision rule (4.1) to object detection

The images shown in Figure 8 demonstrate that the application of the decision rule (4.1) based on the analysis of the cross-correlation coefficient values of frame fragments allows, in many cases, to correctly detect object image fragments. However, there are a small number of fragments that are incorrectly identified as containing the object image.

To reduce the number of incorrectly detected frame fragments in the future, it is advisable to apply the analysis of additional conditions, such as the presence of significant changes in the coordinates of frame fragments corresponding to the maximum values of the cross-correlation coefficient, the degree of change in the values of all calculated coefficients, the average pixel brightness of the analyzed frame fragments, their variance, etc.

5 Conclusion

Thus, in this paper, to evaluate the properties of the sea surface images from the standpoint of solving the problem of detecting floating objects, it is proposed to use the value analysis of the results of calculating the normalized

cross-correlation of video frame sequence fragments. A method is proposed for constructing the analyzed sequence of adjacent frame image fragments and evaluating their differences based on the normalized cross-correlation of the specified fragments.

Computational experiments have shown that the maximum values of the cross-correlation coefficients for a frame fragment sequence containing an object image have values close to 1; when analyzing frame fragments that do not contain an object image, the values of the corresponding coefficients are significantly less than 1. It is shown that the analysis of the proposed correlation properties of the sea surface images allows to develop decision rules for making decisions about the detection of objects.

References

- [1] H. Bay, A. Ess, L.V. Gool, and T. Tuytelaars, *SURF: Speed up robust features*, Comput. Vision Image Understand. (CVIU) **112** (2008), no. 3, 346–359.
- [2] G. Carneiro and D. Lowe, *Sparse flexible models of local features*, Ninth Eur. Conf. Comput. Vision **3** (2006), 29–43.
- [3] V. Golikov and O. Lebedeva, *Adaptive detection of subpixel targets with hypothesis dependent background power*, IEEE Signal Process. Aug. Lett. **21** (2013), no. 8, 751–754.
- [4] V. Golikov, M. Rodriguez-Blanco, and O. Lebedeva, *Robust multipixel matched subspace detection with signal-dependent background power*, J. Appl. Remote Sens. **10** (2016), no. 1, 015006–015006.
- [5] V. Golikov, O. Samovarov, E. Zhilyakov, J.L. Rullan-Lara, and H. Alazki, *Generalized likelihood ratio test for optical subpixel objects' detection with hypothesis-dependent background covariance matrix*, J. Appl. Remote Sens. **14** (2020), no. 4, 046513–046513.
- [6] R.C. Gonzales and P. Wintz, *Digital Image Processing*, Addison-Wesley Longman Publishing Co., Inc., 1987.
- [7] A. Goshtasby, S.H. Gage, and J.F. Bartholic, *A two-stage cross correlation approach to template matching*, IEEE Trans. Pattern Anal. Machine Intell. **6** (1984), no. 3, 374–378.
- [8] R.M. Haralick and L.G. Shapiro, *Computer and Robot Vision*, Addison-Wesley, 1992.
- [9] B.K.P. Horn and B.G. Schunck, *Determining optical flow*, Artif. Intell. **17** (1981), 185–213.
- [10] R. Kerekes and B.V. Kumar, *Enhanced video-based target detection using multi-frame correlation filtering*, IEEE Trans. Aerospace Electronic Sys. **45** (2009), no. 1, 289–307.
- [11] L.L. Scharf and B. Friedlander, *Matched subspace detectors*, IEEE Trans. Signal Process. **42** (1994), no. 8, 2146–2157.
- [12] C. Stauffer and W.E.L. Grimson, *Adaptive background mixture models for real-time tracking*, Proc. IEEE Comput. Soc. Conf. Comput. Vision Pattern Recogn. (Cat. No PR00149), IEEE **2** (1999), 246–252.
- [13] J. Theiler, *Quantitative comparison of quadratic covariance-based anomalous change detectors*, Appl. Opt. **47** (2008), no. 28, F12–F26.
- [14] Y.C. Tzeng, D.M. Chu, M.F. Wu, and C. Kun-Shan, *Automatic detection of targets using fractal dimension*, IEEE Geosci. Remote Sens. Symp. **3** (2005), 1713–1716.
- [15] Z. Wang and J.-H. Xue, *The matched subspace detector with interaction effects*, Pattern Recognition, **68** (2017), 24–37.
- [16] E.G. Zhilyakov, A.A. Chernomorets, E.V. Bolgova, and A.N. Kovalenko, *Image decomposition on the orthogonal basis of subband matrices eigenvectors*, J. Eng. Appl. Sci. **12** (2017), 3194–3197.
- [17] E.G. Zhilyakov, A.A. Chernomorets, E.V. Bolgova, D.V. Ursol, and E.I. Prokhorenko, *On comparing the energy concentration of the image orthogonal transforms*, J. Adva. Res. Dyn. Control Syst. **12** (2020), no. 6, 692–722.
- [18] E.G. Zhilyakov, V. Golikov, D.A. Chernomorets, O.I. Samovarov, and S.L. Babarinov, *Detection of Slow-moving objects floating on an agitated sea surface based on subband analysis within the cosine transform*, J. Adv. Res. Dyn. Control Syst. **12** (2020), no. 25, 1314–1325.