Improving the efficiency of the method of stochastic gradient identification of objects in binary and grayscale images due to their pre-processing

R.G. Magdeev[1,2], M.N. Suetin[1,3], A.G. Tashlinskii[1]

[1]Ulyanovsk State Technical University, Severny Venets str. 32, Ulyanovsk, Russia, 432027
[2]Telekom.ru LLC, Ryabikova str. 21B, Ulyanovsk, Russia, 432045
[3]RPA "Mars", Solnechnaya str. 20, Ulyanovsk, Russia, 432022

Abstract. In this paper, we consider ways to improve the efficiency of the method of stochastic gradient identification of objects for binary and grayscale images due to the methods of image preprocessing. Identification of an object is understood as the recognition of an object on the image with its parameters estimation. Low-pass filtering and image equalization are considered as preliminary processing. The rate of convergence of identification parameters is investigated. The optimal sizes of the Gaussian filter mask for binary and grayscale images were found based on COIL-20 images.

1. Introduction

The task of object recognition, both on separate images and on video sequences, arises in a variety of areas [1, 2]. Establishing the correspondence between the object selected on the image under study and the given patterns (images of object standards) based on a finite set of some properties and signs is the main difficulty. It was shown in [3, 4] that the identification of images of objects by a pattern can be reduced to searching for a spatial transformation that minimizes the metric between the desired image and the pattern. A method of stochastic gradient identification (SGI) of objects in binary images was also proposed there. The method showed good efficiency in comparison with the correlation-extreme method [5] and the contour analysis method [6]. In this paper, we study approaches to increasing the efficiency of SGI for binary images and grayscale images.

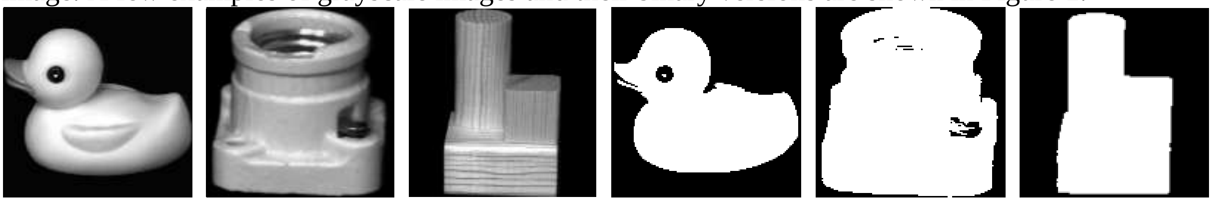
2. Method description

In SGI, the identification parameters $\hat{\vec{\alpha}}$ are searched recursively [7]:

$$\hat{\vec{\alpha}}_t = \hat{\vec{\alpha}}_{t-1} - \mathbf{\Lambda}_t \vec{\beta}_t , \tag{1}$$

where $\vec{\beta}_t$ – stochastic gradient of the objective function of identification quality, which depends on the estimates $\hat{\vec{\alpha}}_{t-1}$ at the previous iteration and on the iteration number $t=\overline{0,T}$; $\mathbf{\Lambda}_t$ – gain matrix [8], which determines the increment of estimates at iterations; T – number of iterations. The stochastic gradient of the objective function at each iteration is calculated from a relatively small (units, tens) local sample of the pattern and the image under study. It is advisable to use the mean square of the inter-frame difference or the inter-frame correlation coefficient as the objective function in the identification problem.

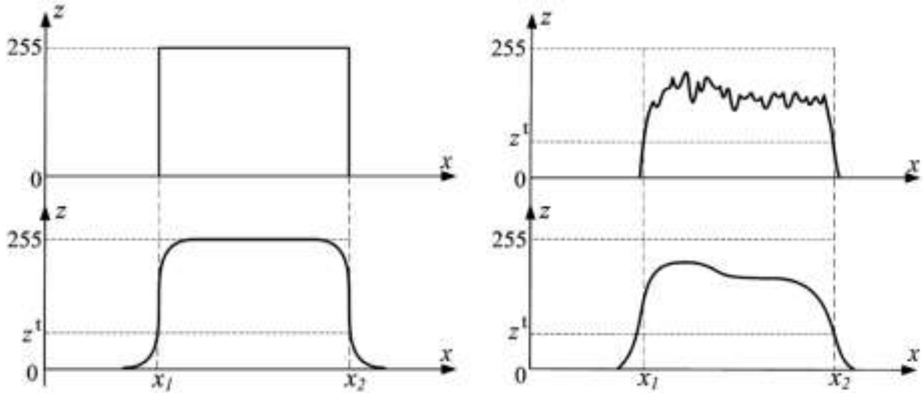
In the study, we suggested that possible deformations of the identified object with respect to the template can be reduced to the model of similarity [9, 10], where the set of parameters $\hat{\bar{\alpha}}$ characterizing the mismatch between the pattern and the image of the object includes a scale factor κ, an orientation angle φ, and a shift $\bar{h} = \left(h_x, h_y\right)^T$ along the base axes Ox and Oy. Independent Gaussian noise was used as additive noise. As the objects of study used the database of images COIL-20 [11], which contains 1440 grayscale images. A binary version was generated for each grayscale image. A few examples of grayscale images and their binary versions are shown in Figure 1.



a) b)

Figure 1. Examples of grayscale patterns (a) and their binary versions (b).

Below are examples of processing the right half-tone and binary images of Figure 1 (wooden bars). In particular, Figure 2 shows graphs of brightness z changes for samples of the 64th row in binary (upper left graph) and grayscale (upper right graph) images, where $z^{\text{п}}$ is the binarization level.



a) b)

Figure 2. The change in brightness along the line, introduced by filtering on binary (a) and grayscale (b) images.

It is not difficult to show that stochastic gradient values $\vec{\beta}_t$ differ from zero only at the points where the brightness derivative of images is also not equal to zero by the evaluated parameters. For a binary image, these are the object boundaries. If at each iteration the local sample is formed randomly (the hit of any image reference is equal to it), the probability of selecting "informational" reference corresponds to the probability of selecting the perimeter of the object:

$$P_{\text{p}} = \mu L_{\text{p}} \Big/ \left(L_x L_y\right) \cong 2\mu / L_x \,, \qquad (2)$$

where: μ - local sample size; L_{p} - the length of the perimeter of the object in samples: L_x and L_y – image size. This estimate is based on the fact that the object image occupies about half the image size. For example, for an image size of 256x256 pixels it is about $0{,}008\mu$. For grayscale images, brightness changes inside the object, so the probability of selecting “informational” counts in the local sample:

$$P_{\text{obj}} \cong 0{,}25\mu \,, \qquad (3)$$

To increase the effective use of SGI on binary images, due to the low probability of choosing “information” samples, it is possible to increase the volume of the local sample and the number of iterations. However, this significantly reduces the speed of the method [10]. In addition, without preliminary processing of binary images, the SGI has a small working range. An example of the convergence of the estimates of the standard deviation (RMS) of the brightness differences between the image and the template (selected as the identification parameter) from the iteration number is shown in Figure 3, where $\hat{\sigma}$ – RMS estimation of the brightness difference. The parameters of the initial shift mismatch were 10 steps of the grid of samples and a rotation of 9 degrees for binary (a) and grayscale (b) images. The figure shows that for the same sample size ($\mu = 21$), the number of iterations required for identification with a binary image and template is approximately twice as large.

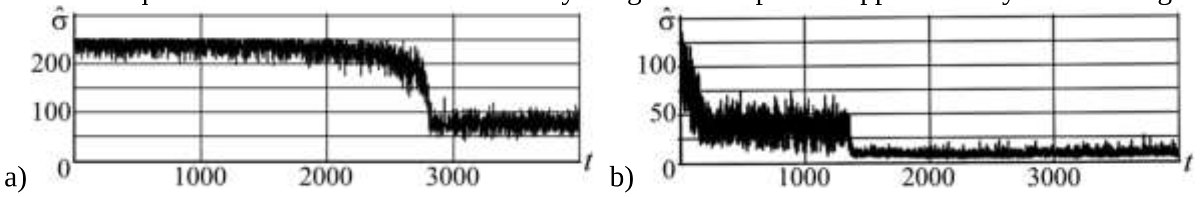


Figure 3. Graphs of estimates of the standard deviation of brightness differences from the number of iterations of stochastic gradient identification for binary (a) and grayscale (b) images.

To improve the efficiency of SGI on binary images, one can artificially increase the number of samples for which the derivative of the objective function is nonzero. For this it is advisable to use "blurring" the boundaries of the object using smoothing filters, for example, a Gaussian filter (lower graphs in Figure 2a). The optimal radius of the filter mask can be determined from the following criteria: rate of convergence of the estimates formed by SGI; maintaining the stability of the shape of the object after the filtering results; working range of identification parameters.

Figure 4a shows the results of experimental studies for objects from the COIL-20 base. At first, the images were artificially subjected to binarization and geometric deformations, after which the parameters of these deformations were estimated using SGI with a preliminary low-pass filtering procedure. The size of the filter mask was selected as a fraction δ of the size of the object in the image. This approach allows you to consider only the ratio of the size of the object and the filter mask and not be tied to the size of the object. The number of iterations T_c to steady state was estimated. In addition, the "distorted" object was compared with another pattern (hereinafter adjacent pattern), which was closest in shape. The averaged correlation coefficient between the "distorted" object and the adjacent template R_{ap} transformed using the SGI was used to assess the stability of the filtering results to the shape of the object. For clarity, in Figure 4, the dotted line also shows the threshold value of identification by the cross-correlation coefficient obtained in [14].

A special way to select samples for a local sample is an additional opportunity to increase the efficiency of SGI when working with binary images. The trick is in the random selection of samples only in the vicinity of the boundary region, and not in the entire image. This allows us not to consider samples in which the derivative of the objective function is a priori zero. Experimental studies have shown [15] that the number of iterations required for identification is reduced by 10-15%.

The use of pre-filtering with the same size of the filter mask for grayscale images can lead to the opposite effect: a decrease in the number of samples in which the derivative of the objective function is nonzero. This can be seen, for example, in Figure 2b (lower graph), where the size of the preliminary filtering mask was 9% of the size of the object. Inside the area of the object there are subregions in which the brightness does not change (they are smoothed out). Therefore, for grayscale images, we conducted a similar study, the results of which are presented in Figure 4b.

The graphs show that the optimal filter mask size for grayscale images is the filter mask size of 2-3% of the image size. Exceeding the threshold for correlation with the adjacent template occurs at 7% of the size of the object. When pre-filtering a grayscale image with a filter mask of 3%, the number of iterations is necessary for the convergence of the estimates $T_{\text{c}} \approx 520$, the average operating range of the SGI when applying such a filter: $\kappa = 0{,}4...1{,}4$; $\varphi = -38^0...+38^0$; $h = -14...+14$ pixels. This is approximately 1.5 times less in the rotation angle parameter and 2.4 times in shift compared to the results obtained in binary images.

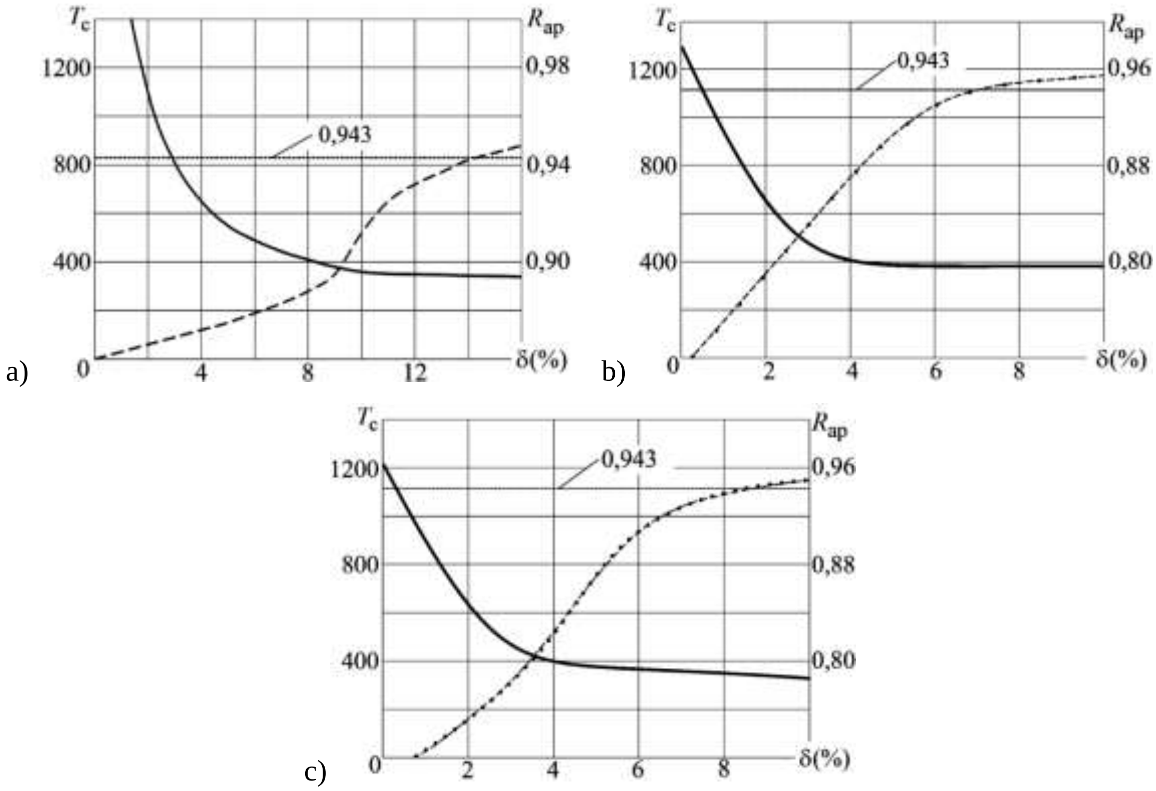


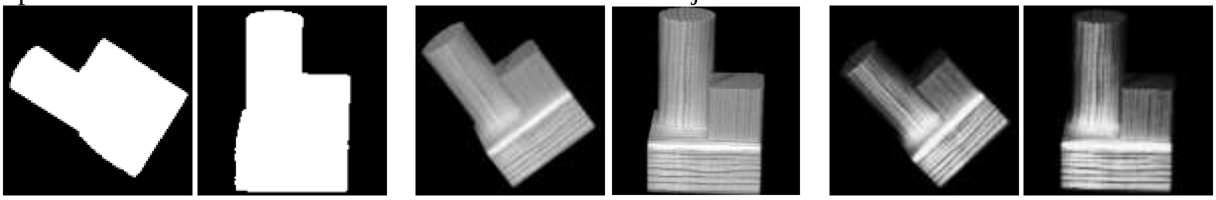
Figure 4. The dependence of the number of iterations to the steady state (continuous curve) and the average correlation with the adjacent pattern (dashed curve) on the size of the Gaussian filter mask (a – for binary images; b – for grayscale images; c – for grayscale images with equalization).

One way to reduce the effect of the size of the filter mask on smoothing gaps inside an object is to pre-equalize it. It allows you to align the histogram of the image and increase the difference in brightness between adjacent pixels [7]. Graphs of the number of iterations to the steady state and the average correlation with the adjacent pattern versus the filter size for grayscale images with preliminary equalization are presented in Figure 4c. The figure shows that the optimal filter mask size for grayscale images with pre-equalization is the filter mask size of 3-4% of the image size of the object, and the threshold for correlation with the adjacent template is exceeded at 9%. The number of iterations necessary for the convergence of estimates during preliminary filtering of a grayscale image with a filter mask of 4%, for which the histogram is preliminary aligned, is $T_c \approx 410$, a the average operating range of SGI is: $\kappa = 0{,}4...1{,}4$; $\varphi = -45^0...+45^0$; $h = -18...+18$ pixels. This is approximately 10% more in terms of the angle and shear parameters compared with the results obtained without equalization.

To illustrate the difference in operating ranges when identifying objects represented by binary (a), grayscale (b) and grayscale with pre-equalization (c) images, Figure 5 shows the identifiable object “wooden bars” with the maximum rotation angle and scale factor that can be estimated SGI. On the right is the template used for all cases.

3. Conclusion

The use of pre-filtering for binary images gives a significant increase in the rate of convergence of estimates formed by the stochastic gradient algorithm. According to the results of studies based on COIL-20 images, the number of iterations to the convergence of identification estimates decreases by almost 10 times (on average from 3600 to 380 iterations) compared with the situation without preliminary filtering. Low-pass filtering has a positive effect on increasing the working range of the stochastic gradient algorithm. The optimal size of the filter mask for binary objects is 9-10% of the size of the object. The results of the study of grayscale images showed that preliminary low-pass filtering of objects is also advisable for them to increase the rate of convergence of identification parameters and expand the effective working range, but due to the peculiarities of grayscale objects, the optimal size of the filter mask is 2-3% of the size of the object. A way to further increase the optimal size of the filter mask is a preliminary equalization procedure, after applying which the optimal size of the filter mask is 3-4% of the size of the object.



a) b) c)

Figure 5. Examples of images of identifiable objects and their patterns.

4. Acknowledgments
The reported study was funded by RFBR & Government of the Ulyanovsk region according to the research projects № 18-41-730006 and № 19-47-730004.

5. References
[1] Poltavskii, A.V. Basics of pattern recognition using computer technology / A.V. Poltavskii, A.V. Grinshkun // Dual technology. – 2017. – Vol. 2. – P. 55-66.
[2] Kniaz, V.A. Intelligent information processing technologies for navigation and control of unmanned aerial vehicles / V.A. Kniaz, B.V. Vishniakov, Iu.V. Vizilter, V.S. Gorbatsevich, O.V. Vygolov // Trudy SPIIRAN. – 2016. – Vol. 2(45). – P. 26-44. DOI: 10.15622/sp.45.2.
[3] Magdeev, R.G. A comparative analysis of the efficiency of the stochastic gradient approach to the identification of objects in binary images / R.G. Magdeev, A.G. Tashlinskii // Pattern recognition and image analysis. – 2014. – Vol. 24(4). – P. 535-541. DOI: 10.1134/S1054661814040130.
[4] Magdeev, R.G. Efficiency of object identification for binary images / R.G. Magdeev, A.G. Tashlinskii1 // Computer Optics. – 2019. – Vol. 43(2). – P. 277-281. DOI: 10.18287/2412-6179-2019-43-2-277-281.
[5] Kovalevskii, V.A. Optimal Image Recognition Techniques / V.A. Kovalevskii – Moscow: Science, 1976. – 328 p.
[6] Furman, Ia.A. Digital methods of processing and recognition of binary images / Ia.A. Furman, A.N. Iurev, V.V. Ianshin – Krasnoiarsk: Izdatelstvo Krasnoiarskogo Universiteta, 1992. – 248 p.
[7] Tashlinskii, A.G. Pseudogradient Estimation of Digital Images Interframe Geometrical Deformations / A.G. Tashlinskii // Vision Systems: Segmentation & Pattern Recognition, 2007. – P. 465-494. DOI: 10.5772/4975.
[8] Tsypkin, Ya.Z. Information theory of identification / Ya.Z. Tsypkin – Moscow: Fizmatlit, 1995. – 336 p.
[9] Gonzalez, R. Digital image processing / R. Gonzalez, R. Woods // Upper Saddle River: New Jersey: Prentice Hall, 2012. – 1012 p.
[10] Fadeeva, G.L. Optimization of the pseudogradient of the objective function when evaluating interframe geometric deformations of images: Disertatsiia kandidata tekhnicheskikh nauk: 05.13.18. – Ulianovsk, 2000. – 17 p.
[11] Columbia Object Image Library (COIL-20) [Electronic resource]. – Access mode: http://www.cs.columbia.edu/CAVE/software/softlib/coil-20.php (24.12.2019).
[12] Tashlinskii, A.G. The specifics of pseudogradient estimation of geometric deformations in image sequences / A.G. Tashlinskii // Pattern Recognition and Image Analysis. – 2008. – Vol. 18(4). – P. 700-705. DOI: 10.1134/S1054661808040275.
[13] Tashlinskii, A.G. Estimation of spatial deformation parameters of image sequences / A.G. Tashlinskii – Ulyanovsk: ULSTU, 2000. – 131 p.

[14] Tashlinskii, A.G. Improving the reliability of identification of objects in binary images / A.G. Tashlinskii, R.G. Magdeev // Information-measuring and control systems. – 2017. – Vol. 15(12). – P. 24-30.
[15] Magdeev, R.G. Estimation of the probability of disruption during pseudo-gradient search for image fragments / R.G. Magdeev, A.G. Tashlinskii, P.V. Iakshankin // Radio-electronic devices and systems for infocommunication technologies, 2013. – P. 324-326.