Efficiency of Stochastic Gradient Identification of Similar Shape Objects in Binary and Grayscale Images

Radik Magdeev
LLC "Telecom.ru"
Ulyanovsk, Russia
radiktkd2@yandex.ru

Aleksander Tashlinsky
Radio Engineering Department
Ulyanovsk State Technical University
Ulyanovsk, Russia
tag@ulstu.ru

Galina Safina
National Research Moscow State University of Civil Engineering
Moscow, Russia
safinagl@mgsu.ru

***Abstract*—A comparative analysis of efficiency of stochastic gradient identification method on the base of pattern of objects with similar shapes by their grayscale and binary images is carried out. Object identification is understood as the determination of the object image in the studied image with the estimation of its spatial parameters in relation to the reference image. Two types of objects with similar shape are investigated on the base of COIL-20 halftone images and their binary versions. The objects of the first type have a different character of the curvature of the lines describing their contour, and the objects of the second type are close to the curvature characteristics of the contour lines.**

***Keywords*—binary image, grayscale image, object recognition, pattern recognition, stochastic gradient identification, parameter estimation, convergence**

I. Introduction

The problem of pattern recognition, both on separate images and on video sequences, arises in a variety of areas: from military affairs and security systems to the digitization of analog signals. The problem of automating the solution of this problem remains relevant both from the point of view of theory and technical implementation [1-3]. Pattern recognition, as a rule, is considered as assigning on the basis of the initial data of the object in the image, to a certain class (group of classes) by comparing the selected essential features characterizing this class. The main difficulty in this case is to establish the correspondence between the object highlighted in the studied image and the given patterns (images of the object's standards) based on a finite set of some properties and attributes. Note that there are several areas in pattern recognition:

– recognition of many predefined objects, or classes of objects in the image;

– object detection, implemented by checking the image or its part for compliance with certain conditions;

– identification on the image of the object with the assessment of its parameters and decision making.

In [4, 5] it is shown that identifying images of objects by a pattern can be reduced to searching for a spatial transformation that minimizes the distance between the desired image and the pattern in a given metric space, and a stochastic gradient identification method (SGIM) of objects on binary images is proposed, which showed good efficiency in comparison with the correlation-extreme method [6] and the contour analysis method [7]. This article discusses the effectiveness of SGIM for grayscale images in comparison with its usage for binarized images.

For concreteness, we will assume that possible deformations of the identified object with respect to the pattern can be reduced to a similarity model [8, 9], that is, the pattern and image of the object can differ in scale factor κ, orientation angle φ, and shifts $\bar{h} = (h_x, h_y)^T$ along the basic axes Ox and Oy, in addition, additive noise. We used the COIL-20 halftone images including images of 1440 objects [10]. In this case, binary versions were obtained for each of the halftone images. A number of examples of halftone images and their binary versions are shown in Fig. 1.

II. Identification Method Description

In SGIM the identification parameters $\vec{\alpha}$, on the basis of which the decision is made, are searched recursively [11]:

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

where $\vec{\beta}_t$ is the stochastic gradient of the cost function of identification quality, depending on $\hat{\vec{\alpha}}_{t-1}$ and the iteration number $t = \overline{0,T}$; $\mathbf{\Lambda}_t$ is the gain matrix [12]; T is the number of iterations. It was shown in [11, 13] that it is advisable to use the brightness correlation coefficient (BCC) or the mean square of the brightness difference (MSBD) of the pattern and the studied image as the cost function, which were used in this work. Hereinafter, a pattern refers to a reference image of an object. At each iteration, in order to find the next estimate of the parameter vector two-dimensional local sample of the same samples on the pattern and the studied image is used. As a rule, this sample has small size [14].

The effective working range of the estimated parameters of the SGIM (in which the estimates for a given number of iterations do not go beyond the required confidence interval) is limited. If it does not cover the domain of parameters, then to provide coverage it is required to specify several patterns with different initial approximations of the parameters. It was also shown in [4, 5] that in order to increase the convergence rate of estimates and to expand the working range for binary images it is advisable to use low-pass filtering, for example, Gaussian, as the pre-processing. The optimal size of the mask of a Gaussian filter for binary images is 10 % of the identified object size.

It was also shown in [4, 5] that in order to increase the convergence rate of estimates and to expand the working range for binary images, it is advisable to use low-pass filtering, for example, Gaussian, as the pre-processing. The optimal size of the mask of a Gaussian filter for binary images is 10% of the identified object size.

The studies using halftone images from the COIL-20 base have also shown the appropriateness of low-pass filtering. In this case, the optimal size of the Gaussian filter mask, which allows expanding the operating range of the SGIM while maintaining identification accuracy, is from 3 % to 10 % of the of the object size in the image. We also note that the approximate implementation of the Gaussian filter proposed in [15, 16] and based on infinite impulse response is used. The computational complexity of the approach used does not depend on the size of the filter mask and is approximately $16 L_x L_y$ elementary operations, where L_x and L_y are the image sizes.

(a)

(b)

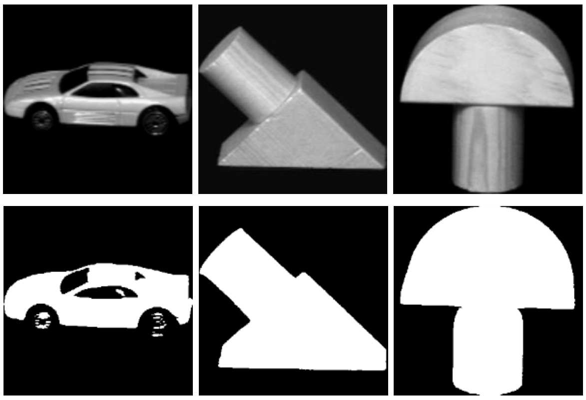


Fig. 1. Example of halftone patterns (a) and their binary versions (b).

The computational complexity of the stochastic gradient parameter estimation procedure that underlies the SGIM was studied in [17] and, in particular, is similar to the parameters of the similarity model when using MSBD from $(22\mu + 25)T$ to $(52\mu + 20)T$ elementary operations (depending on the chosen method of finding the pseudo-gradient of the objective function), and when using the BCC from $(51\mu + 91)T$ to $(69\mu + 48)T$ elementary operations, where μ is the local sample size at each iteration.

As a characteristic of the SGIM efficiency for binary and grayscale images, we use the convergence of the standard deviation (SD) $\hat{\sigma}_t$ of the brightness differences of the modified pattern and the studied image, which is calculated at each t -th iteration from a local sample of identifiable image and pattern samples, $t = \overline{0,T}$. Example of $\hat{\sigma}_t$ convergence graphs for the left object of Fig. 1 (car) with the mismatch parameters of the pattern and the studied object: $\kappa = 0.85$, $\varphi = 35^0$, $\bar{h} = \left(h_x, h_y\right)^T = (6, -6)^T$, is shown in Fig. 2, where graph (a) corresponds to a halftone image, and (b) a binary image. The studied images and corresponding patterns are shown in Fig. 3, and the convergence graphs of the estimates of individual identification parameters are shown in Fig. 4, where the solid line corresponds to the grayscale images and the dashed line corresponds to binary images. The image sizes are 128x128 elements, the local sample size is $\mu = 15$.

It can be seen from the plots that for this object estimates of the identification parameters $\vec{\hat{\alpha}}_t = \left(\hat{\kappa}_t, \hat{\varphi}_t, \vec{\hat{h}}_t\right)^T$ when processing halftone images and patterns converge slower (for about 400 iterations) than when processing their binarized versions (for about 200 iterations). This is explained by the large size of the low-pass filter during image preprocessing. Thus, both the rate of convergence of estimates and the effective working range when using grayscale and binarized images can vary. This is especially true for images of objects having a similar shape. $\hat{h}_{xt}$

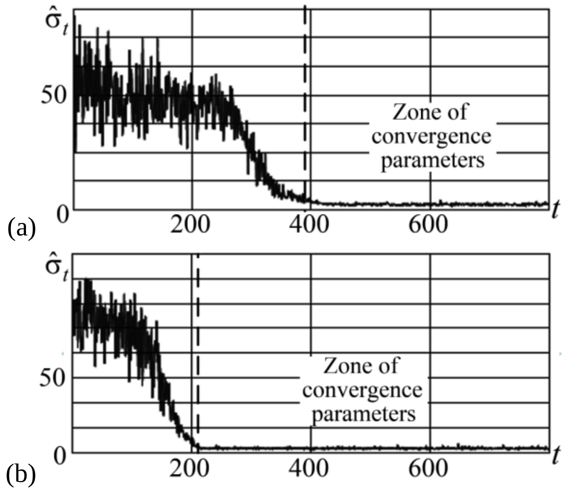


Fig. 2. Convergence of brightness differences SD of the modified pattern and the studied image for grayscale (a) and binary (b) images.

(a)

(b)

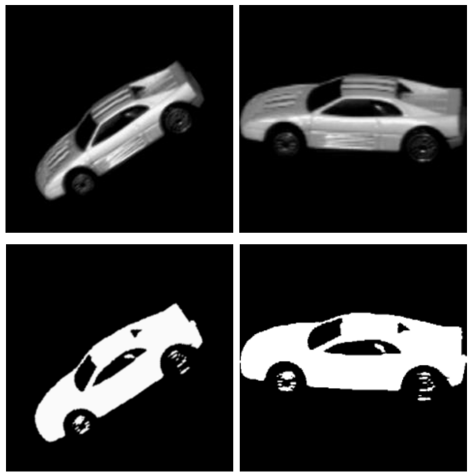


Fig. 3. Example of studied halftone (a) and binary (b) images and their corresponding patterns.

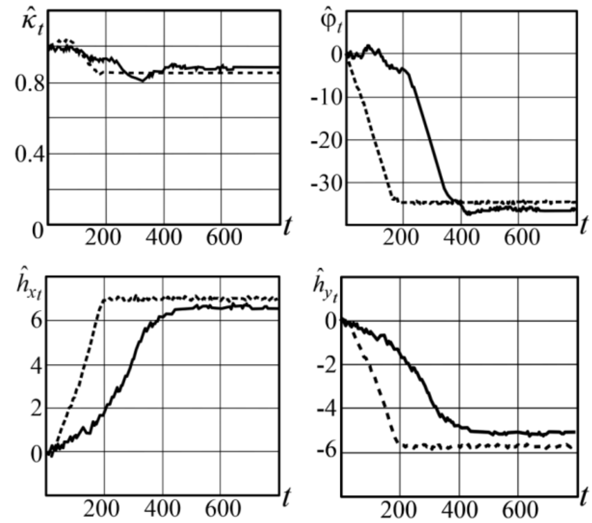
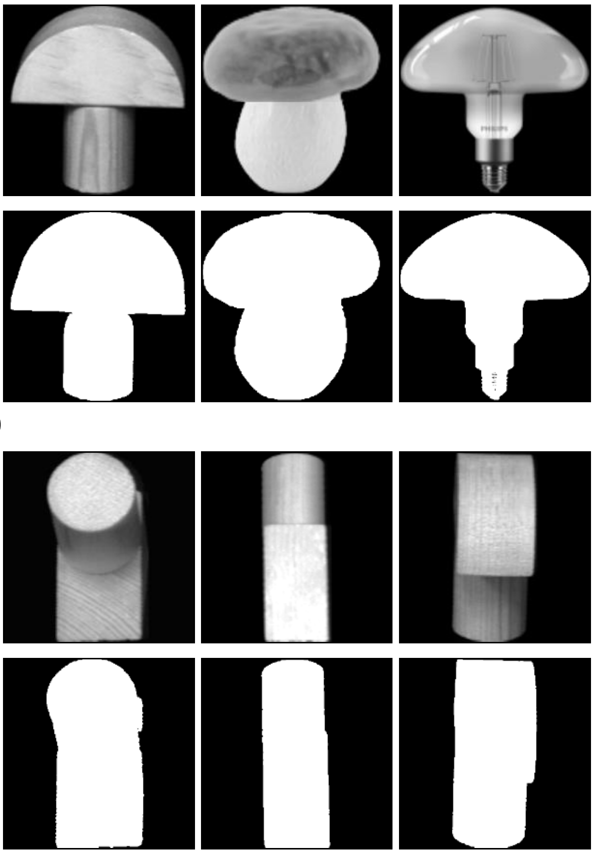


Fig. 4. Iidentification parameters convergence.

III. Identification of Objects With a Similar Shape

Using the objects of the COIL-20 images, we consider two types of objects that are similar in shape: the curvatures of the lines describing the contour of a different nature (the objects shown in Fig. 5a can serve as an example), and with similar curvature characteristics of the contour lines (an example of such objects is shown in Fig. 5b). The indicated figures also show binary versions of the images of these objects. Obviously, the studied types of objects are critical in the processing of binary images.



(a)

(b)

Fig. 5. Examples of similarly shaped images having different and close characteristics of the contour lines curvature.

In the experiment, the identification method proposed in [18] was applied and based on three criteria, one of which uses the correlation coefficient between the studied (deformed) image of the object and the patterns transformed using SGIM (we will conventionally call this criterion the main one). Two other criteria use convergence characteristics of identification parameters (additional criteria). One characteristic is the estimation of the mean value of the standard deviation of the brightness differences of the modified pattern and the studied image in the steady state of the process of evaluating the SGIM identification parameters. This characteristic is in iterations of steady state. Another characteristic is the standard deviation of values, also at iterations of the steady state.

The steady state of the identification process is clearly illustrated in Fig. 2. The decision on identification is made if all three criteria are fulfilled:

$$R \geq R^t,\ m_{\hat{\sigma}_t} \leq m^t_{\hat{\sigma}_i},\ \delta_{\hat{\sigma}_i} \leq \delta^t_{\hat{\sigma}_i},$$

where R^t, $m^t_{\hat{\sigma}_i}$ and $\delta^t_{\hat{\sigma}_i}$ are threshold values. The threshold values of the identification criteria for the used image database were determined by the method [18]:

$$R^t = 0.92,\ m^t_{\hat{\sigma}_i} = 9.16,\ \delta^t_{\hat{\sigma}_i} = 4.63.$$

The following results are obtained for the first type of objects. For the binarized images, the correlation coefficient between the image of the object and the “correct” pattern is $R = 0.99$ and exceeds the threshold value. For this pair the additional criteria are also fulfilled:

$$m_{\hat{\sigma}_t} = 1.11 < m^t_{\hat{\sigma}_i},\ \delta_{\hat{\sigma}_i} = 0.69 < \delta^t_{\hat{\sigma}_i}.$$

However, the correlation coefficient between the image of the object and similar patterns transformed by the SGIM also exceeds the identification threshold value ($R \approx 0{,}94$). At the same time, the numerical values of auxiliary characteristics do not reach threshold values, although they are quite close to them ($m_{\hat{\sigma}_t} \approx 11, \delta_{\hat{\sigma}_i} \approx 7.3$). For grayscale images, the correlation coefficient between the image of the object and the “correct” pattern is also 0.99 and exceeds the threshold value, and the correlation coefficient with similar patterns ($R \approx 0.7$) is significantly lower than the threshold. Additional criteria for the “correct” pattern are also fulfilled:

$$m_{\hat{\sigma}_t} = 1.21,\ \delta_{\hat{\sigma}_i} = 0.27$$

and for similar pattern the values of additional characteristics significantly exceed the threshold:

$$m_{\hat{\sigma}_t} \approx 17 > 9.61, \delta_{\hat{\sigma}_i} \approx 15 > 4.63.$$

Thus, for this type of object, when binarizing their images, the decision on identification requires the use of additional criteria. For grayscale images, a decision on identification is possible using only the main criterion for the correlation coefficient, and additional ones can be used to assess the reliability of the identification.

An analysis of the usage of SGIM for binary images of objects of similar shape in the second type showed that all identification criteria are satisfied, both for the “correct” pattern and for similar ones. So, for the “correct” pattern are:

$$R = 0.99,\ m_{\hat{\sigma}_t} = 1.31 < m^t_{\hat{\sigma}_i},\ \delta_{\hat{\sigma}_i} = 0.89 < \delta^t_{\hat{\sigma}_i}$$

and for similar are:

$$R \approx 0{,}97,\ m_{\hat{\sigma}_t} \approx 7.2 < m^t_{\hat{\sigma}_i},\ \delta_{\hat{\sigma}_i} \approx 1.4 < \delta^t_{\hat{\sigma}_i}.$$

For grayscale images, the correlation coefficient between the image of the object and the “correct” pattern exceeds the threshold, but less than in the other cases considered ($R = 0.96$). The values of the additional characteristics are significantly lower than the threshold:

$$m_{\hat{\sigma}_t} = 1.81,\ \delta_{\hat{\sigma}_i} = 1.74.$$

For similar pattern, the criterion for the correlation coefficient is not satisfied ($R \approx 0.83$) and the values of the auxiliary characteristics significantly exceed the threshold ($m_{\hat{\sigma}_t} \approx 23.3$, $\delta_{\hat{\sigma}_i} \approx 12.3$).

Thus, for objects of similar shape with similar characteristics of contour lines curvature, their identification by the pattern from binarized images is ineffective. When identifying this type of objects by their grayscale images, it is advisable to use the additional criteria used in the work.

It should be noted that the effective operating range of SGIM for images of this type is significantly reduced. So, when choosing similarity model parameters as identification parameters, for the images considered, it is: $\kappa = 0.8 \dots 1.1$; $\varphi = -10^0 \dots +10^0$; $h = -5 \dots +5$ pixels. This is due to the fact that such images differ mainly in texture, and the preliminary low-pass filtering procedure smooths the texture, so the size of the preprocessing filter mask does not exceed 3 % of the size of the object.

IV. Conclusion

A comparative analysis showed that the extension of the SGIM to grayscale images does not impair its performance in computational complexity, but slightly reduces the effective operating range and the convergence rate of identification parameters. This is due to the fact that with the same reliability of identification of objects, grayscale images allow a smaller size of the low-pass filter during pre-processing.

A study on the basis of COIL-20 images of identifying objects of similar shape with different curvatures characterizing the lines of an object's contour showed that with close characteristics of the curvature of contour lines, identification by binarized images is ineffective. When identifying by grayscale images, it is advisable to increase the reliability of using, in addition to correlation criteria, additional ones based on the characteristics of the process of convergence of identification parameters. For objects of similar shape with different characteristics of the contour lines curvature, when binarizing their images, the decision to identify also requires the use of additional criteria. For grayscale images, a solution is possible using only the correlation criterion, and additional ones can be used to assess the reliability of identification.

We also note that in order to solve the problem of identification of objects according to a pattern the criterion based on correlation and referred in this paper as "basic" is not significant for identification in many cases.

Acknowledgment

The reported study was funded by RFBR & Government of the Ulyanovsk region according to the research projects № 19-29-09048 and № 19-47-730004.

References

[1] S.M. Borzov, M.A. Guryanov and O.I. Potaturkin, "Study of the classification efficiency of difficult-to-distinguish vegetation types using hyperspectral data," Computer Optics, vol. 43, no. 3, pp. 464-473, 2019. DOI: 10.18287/2412-6179-2019-43-3-464-473.

[2] A.V. Poltavskii and A.V. Grinshkun, "Basics of pattern recognition using computer," Dvoinie Tehnologii, vol. 2, pp. 55-66, 2017

[3] S.V. Kurochkin, "Detection of the homotopy type of an object using differential invariants of an approximating map," Computer Optics, vol. 43, no. 4, pp. 611-617, 2019. DOI: 10.18287/2412-6179-2019-43-4-611-617.

[4] R.G. Magdeev and A.G. Tashlinskii, "A comparative analysis of the efficiency of the stochastic gradient approach to the identification of objects in binary images," Pat. Rec. and Im. Anal., vol. 24, no. 4, pp. 535-541, 2014. DOI: 10.1134/S1054661814040130.

[5] R.G. Magdeev and A.G. Tashlinskii, "Efficiency of object identification for binary," Computer Optics, vol. 43, no. 2, pp. 277-281, 2019. DOI: 10.18287/2412-6179-2019-43-2-277-281.

[6] V.A. Kovalevsky, "Methods of optimal solutions in image recognition," Moscow: Nauka Publisher, 2003.

[7] Ya.A. Furman, A.V. Krevetsky, A.K. Peredeyev, A.A. Rozhentsov, R.G. Khafizov, I.L. Egoshina and A.L. Leukhin, "Introduction to contour analysis and its applications to image and signal processing," Moscow: Fizmatlit Publisher, 2003.

[8] R. Gonzalez and R Woods, "Digital image processing," Upper Saddle River: New Jersey: Prentice Hall, 2012.

[9] A.G. Tashlinskii, "Estimation of the parameters of spatial deformations of image sequences," Ulyanovsk: UlSTU Publisher, 2000.

[10] S.A. Nene, S.K. Nayar and Y. Murase, "Columbia Object Image Library," COIL-20 [Online]. URL: http://www.cs.columbia.edu/CAVE/software/softlib/coil-20.php.

[11] A.G. Tashlinskii, "Pseudogradient Estimation of Digital Images Interframe Geometrical Deformations," Vision Systems: Segmentation & Pattern Recognition, pp. 465-494, 2007. DOI: 10.5772/4975.

[12] Ya.Z. Tsypkin, "Information theory of identification," Moscow: Fizmatlit Publisher, 1995.

[13] A.G. Tashlinskii, "The specifics of pseudogradient estimation of geometric deformations in image sequences," Pattern Recognition and Image Analysis, vol. 18, no. 4, pp. 706-711, 2008. DOI: 10.1134/S1054661808040287.

[14] A.G. Tashlinskii, G.V. Dikarina, G.L. Minkina and A.N. Repin, "Pseudogradient optimization in the estimation of geometric interframe image deformations," Pattern recognition and image analysis, vol. 18, no. 4, pp. 535-541, 2008. DOI: 10.1134/S1054661814040130.

[15] L.J. van Vliet, I.T. Young and P.W. Verbeek, "Van Vliet Recursive Gaussian Derivative Filters," Proc. 14th Int. Conference on Pattern Recognition, pp. 509-514, 1998. DOI: 10.1109/ICPR.1998.711192.

[16] D. Hale, "RecursiveGaussianfilters" [Online]. URL: https://inside.mines.edu/~dhale/papers/Hale06RecursiveGaussianFilters.pdf.

[17] G.L. Fadeeva, "Optimization of the pseudo-gradient of the objective function in the estimation of inter-frame geometric deformations of images," dis. cand. tech. Sciences: 05.13.18, Ulyanovsk, 2007.

[18] A.G. Tashlinskii and R.G. Magdeev, "Increasing objects identification accuracy for binary images," Informatsionno-izmeritel'nyye i upravlyayushchiye sistemy, vol. 12, pp. 24-30, 2017.