

Efficiency of object identification for binary images

R.G. Magdeev^{1,2}, A.G. Tashlinskii¹

¹Ulyanovsk State Technical University, Russia, Ulyanovsk,

²Telekom.ru LLC, Russia, Ulyanovsk

Abstract

In this paper, a comparative analysis of the correlation-extreme method, the method of contour analysis and the method of stochastic gradient identification in the objects identification for a binary image is carried out. The results are obtained for a situation where possible deformations of an identified object with respect to a pattern can be reduced to a similarity model, that is, the pattern and the object may differ in scale, orientation angle, shift along the base axes, and additive noise. The identification of an object is understood as the recognition of its image with an estimate of the strain parameters relative to the template.

Keywords: digital image, object recognition, pattern recognition, correlation-extreme algorithm, stochastic gradient identification, incorrect identification probability.

Citation: Magdeev RG, Tashlinskii AG. Efficiency of object identification for binary images. *Computer Optics* 2019; 43(2): 277-281. DOI: 10.18287/2412-6179-2019-43-2-277-281.

Acknowledgments: This work was supported by RFBR and the government of Ulyanovsk region, project no. 16-47-732053 and the RFBR grant, project no. 18-41-730006.

Introduction

The problem of pattern recognition, both for images and for video sequences, occurs in a variety of areas: from military affairs and security systems to the digitisation of analog signals. Automation of the solution of this problem is still a complex problem both from the theoretical and technical point of view [1–3]. In this case, the recognition of images can be considered as a reference on the basis of the original data of the object in the image to a certain class (group of classes) by comparing the distinguished essential characteristics that characterise a given class. The main problem in this case is to establish a correspondence between the object selected on the image and the given template based on the ratio of the final set of certain properties and characteristics. There are the following main approaches in pattern recognition:

- recognition of a set of predefined objects or classes of objects in the image;
- object detection, consisting in checking the image or its part for compliance with certain conditions;
- identification of the object on the image with an estimation of its parameters and decision making.

It is shown in [4] that the problem of identifying images of objects with a template can be reduced to the search for a spatial transformation, which minimises the distance between the target image and the template in a given metric space.

This work performs a comparative analysis of correlation-extreme (CEM) [5, 6] method, contour analysis method (CAM) [7], and stochastic gradient identification algorithm (SGI) [8, 9] in conditions of similarity model [10] deformation between referenced and target objects, i.e. they can differ by translation $\vec{h} = (h_x, h_y)^T$ along base axis Ox and Oy , rotation ϕ and scale κ in addition to additive noise.

CAM and SGI work directly with the image of objects or geometrical signs of objects on the image (contours). CEM works in both spatial and frequency domains.

Comparison of the selected methods is based on computational complexity and the probability of false identification of the object.

1. Computational complexity

Let us estimate the computational complexity as the number of elementary mathematical and logical operations of the implementations of the methods analyzed.

1.1. Correlation-extreme method

The idea of the CEA can be reduced to a computation of the normalized correlation function of the sought image and the pattern image for all specified possible conversion parameters [6]. If there is a similar fragment in the initial image, the maximum of the correlation function arises from this part. The basic stages of the CEA are computation of the correlation coefficient for all possible positions of the object (with all patterns), discovery of the maximal coefficient, and its comparison with the threshold, ensuring the specified probability of correct identification. The computational complexity of the CEA depends on the definition region of possible parameter values and, when the size of the pattern image is $w \times l$ elements, it approximates to:

$$C_{CEA}^D \approx 4k_\kappa k_\phi k_{h_x} k_{h_y} (wl + 1),$$

where: $k_{h_x} = (W - w) / \Delta h$, $k_{h_y} = (L - l) / \Delta h$, $k_\kappa = (\kappa_{\max} - \kappa_{\min}) / \Delta \kappa$ и $k_\phi = ((\phi_{\max} - \phi_{\min}) / \Delta \phi)$ – the number of similar templates for the defined parameters \vec{h} , κ and ϕ respectively; $\phi_{\max(\min)}$ and $\kappa_{\max(\min)}$ – maximal (minimal) rotation angle and scale factor; $\Delta \kappa$, $\Delta \phi$ and Δh – increment of corresponding parameter change; $W \times L$ – the size of the studied image. If the orientation of the object is not limited, we obtain:

$$C_{CEA}^D \approx \frac{8\pi(\kappa_{\max} - \kappa_{\min})(W - w)(L - l)(wl + 1)}{(\Delta h)^2 \Delta \kappa \Delta \phi}.$$

Some decrease in the computational burden allows a transition to the frequency region in some cases. The

transition is carried out in accordance with the discrete Fourier transform. The fast Fourier transform with a computational complexity of $WL \log(WL)$ ensures a higher speed of operation. The study of the amplitude-frequency characteristics of pattern images allows one to virtually exclude the computational burden related to finding the parameters of shift \vec{h} . Thus, the computational complexity of the CEA in the frequency region is:

$$C_{CEA}^F \approx \frac{2\pi(\kappa_{\max} - \kappa_{\min})(WL)(\log(WL) + 4)}{\Delta\kappa \Delta\phi}$$

1.2. Contour analysis method

The CAM allows one to recognize objects, represented by their external outlines, i.e., contours. To extract information on the shape of the object, the contour is specified as a closed-vector contour [7]. The length of the contour (the number ℓ of its components' elementary vectors), encoded with a two-dimensional code, is normalized. Then, a normalized correlation function of the obtained contour vector and the vector formed from the pattern by cyclically shifting its elementary vectors (specifying the mutual shift of the contours), is calculated. The excess of the correlation function module over the preset threshold corresponds to the identification of the object. The main stages of the CAM assuming the use of the Canny approach to the edge detection of objects [11], in which the Gaussian filters and the fast Fourier transform [10] are applied for noise suppression, and the Sobel operator [12] does for finding gradients and the number of elementary operations required for their implementation are summarized in table 1.

Thus, the computational complexity of the Canny approach is as follows:

$$C_{CAM} \approx 2WL(\log(WL) + 15) + 16(w + l) + 6\ell^2 + 4\ell$$

Table 1. Computational complexity of the CAM

	Number of operations
Noise suppression	$2WL \log(WL)$
Search for gradients	$12WL$
Suppression of local boundary maxima in the gradient direction	$8WL$
Search for gradients	$2WL$
Double threshold filtering	$8WL$
Route location of the ambiguity zone	$2WL \log(WL)$
Representation of contours in the vector form	$16(w + l)$
Normalization of the contour length	4ℓ
Calculation of the normalized correlation function	$4\ell^2$

1.3. Stochastic gradient identification algorithm

In the SGI algorithm the identification parameters are searched recurrently [13]:

$$\hat{\alpha}_t = \hat{\alpha}_{t-1} - \Lambda_t \bar{\beta}_t,$$

where $\bar{\beta}_t$ – stochastic gradient of the objective function Q , depending on $\hat{\alpha}_{t-1}$ and the iteration number $t = \overline{0, T}$; Λ_t – amplification matrix [6]. In the identification prob-

lem, the coefficient of interframe correlation coefficient (CC) is often selected as Q [14].

The working range of the estimated parameters (where the estimates do not overstep the limits of the required confidence interval and when the number of iterations is specified) of the SGI algorithm is limited. If it does not cover the parameter definition region, several patterns must be specified with different initial approximations of parameters to ensure coverage. To increase the speed of estimate convergence $\bar{\alpha}$ and to expand the working range of the SGI algorithm, it is expedient to apply low-frequency, e.g. Gaussian, filtering to binary images. As already stated, for this purpose, approximately $2WL \log(WL)$ elementary operations are required. The computational complexity of the SGI algorithm is considered in [15] and when mean squared error (MSE) is chosen as Q it is ranges from $(22\mu + 25)T$ to $(52\mu + 20)T$ elementary operations, and for CC - between $(51\mu + 91)T$ and $(69\mu + 48)T$ elementary operations, where μ – sample size on each iteration, T – the number of iterations.

Hence, computational complexity of SGI with MSE in average can be found as follows:

$$C_{SGI}^{MSE} \approx 2MN(\log(MN) + 15) + (32\mu + 24)T$$

for CC:

$$C_{SGI}^{CC} \approx 2MN(\log(MN) + 15) + (60\mu + 70)T,$$

Figure 1a shows the dependences of the computational complexity of the studied methods on the image size $W=L$ with a constant object size of $w=l=128$ elements. The curve 1 corresponds to the CEA in the spatial region at $\kappa_{\max}=1.4$, $\kappa_{\min}=0.6$, $\Delta h=2$, $\Delta\kappa=0.2$, $\Delta\phi=0.2$; curve 2 – CEA in the frequency region at the same parameter values; curve 3 – CAM with $\ell=50$; curves 4 and 5 – SGI with $\mu=20$, $T=2000$ and MSE or CC as Q respectively (similar notations for the curves were used in the remaining figures). One can see that if the image contains below $5 \cdot 10^5$ pixels, the CAM has a lower computational complexity. The computational complexity of the CEA in the spatial region is higher by about two orders of magnitude and in an approximately quadratic relation to the image size which is weakly expressed in the figure. The computational complexity of the CEA in the frequency region substantially depends on image size. When $W \approx 500$ it is lower by an order of magnitude than for the spatial region and, when $W \approx 3500$ it is higher by an order of magnitude.

Figure 1b shows the dependences of computational complexity on the object size $w=l$ at a constant image size $W=L=1024$ and with the same methodical characteristics. One can see that the computational complexity of CAM, SGI, and CEA depends weakly on object size in the frequency region, and that for the CEA it is approximately quadratic in the spatial region. The SGI with the MSFD requires a minimum computational cost, and the CEA requires a maximum computational cost.

The experiment, performed with the use of a computer with an AMD Athlon II X2 250 processor with a

3.00-GHz frequency at $W=L=512$, $w=l=256$ and 200 samples, had an CEA average operation time in the spatial region of about 18 min and in the frequency region of 2 min, whereas the CAM was 0.6 s, SGI (MSE) was 0.78 s, and SGI (CC) was 0.92 s. It is necessary to note that three initial angle approximations were specified for the SGI algorithm, since the working range of the method with the iteration number used is about $\pm 60^\circ$. A calculation of computational complexity for the same values gave $C_{CEA}^D \approx 1,5 \cdot 10^{10}$, $C_{CEA}^F \approx 4,6 \cdot 10^9$, $C_{CAM} \approx 1.1 \cdot 10^7$, $C_{SGI}^{MSE} = 1,6 \cdot 10^7$, $C_{SGI}^{CC} = 1,7 \cdot 10^7$ which was in agreement with experimental data.

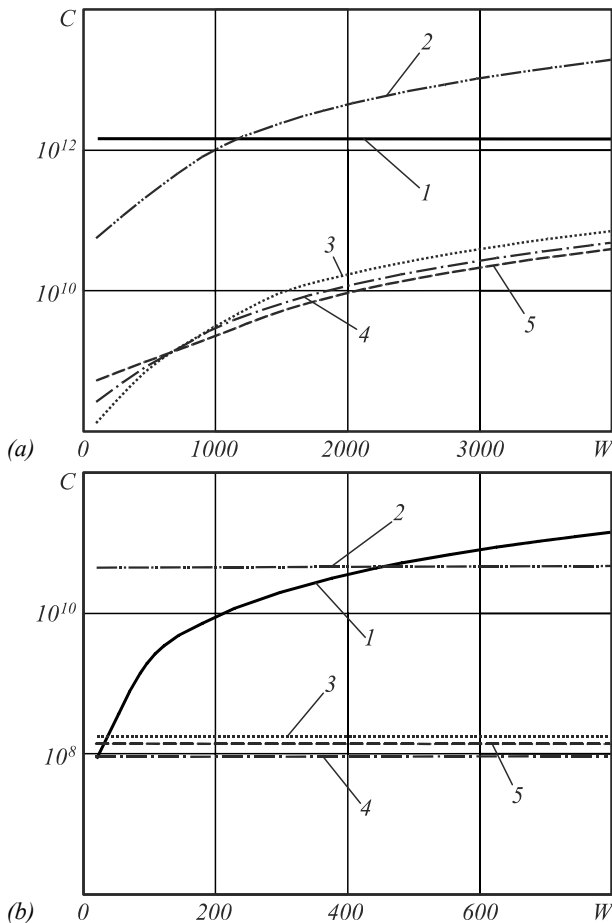


Fig. 1. Dependence of the computational complexity of the methods on image size (a) and object size (b)

2. Probability of false identification

The probability of false identification P_{er} was determined experimentally. In this case, the influence of additive noise was studied in the range of signal/noise ratio (q) for dispersions of 1-10 and mismatch of location of initial and pattern objects, which is critical for SGI. The dependences of P_{er} on the signal/noise ratio are given in figure 2a. The CEA showed the best noise immunity in the spatial region due to its large sample volume. Here the erroneous identification is basically caused by a sufficiently large identification parameter step between the patterns ($\Delta h=5$, $\Delta \kappa=0.2$, $\Delta \phi=8^\circ$), which it was difficult to decrease in the experiment due to the large computational cost. If the object sought has a high-frequency spa-

tial spectrum, the noise immunity of the CEA in the frequency region is much worse than in the spatial region. Note that the PGIA ensures a high noise immunity, and it gave the least P_{er} at small levels of noise ($q > 8$). This can be attributed to the high identification accuracy of the location parameters of the object sought. The noise immunity of the CAM is several times weaker due to errors in the detection of contours over the whole range q .

Figure 2b shows the dependences of the probability of false identification on the location mismatch of the pattern and the object at $q=10$. One can see that this parameter is critical only for the SGI, which has a limiting working range. Here, when the mismatch changes from 0 to 40 sample grid intervals, P_{er} increases by factor of approximately 4, when the CC is selected as Q , and by factor of 5, when the MSE is selected as Q .

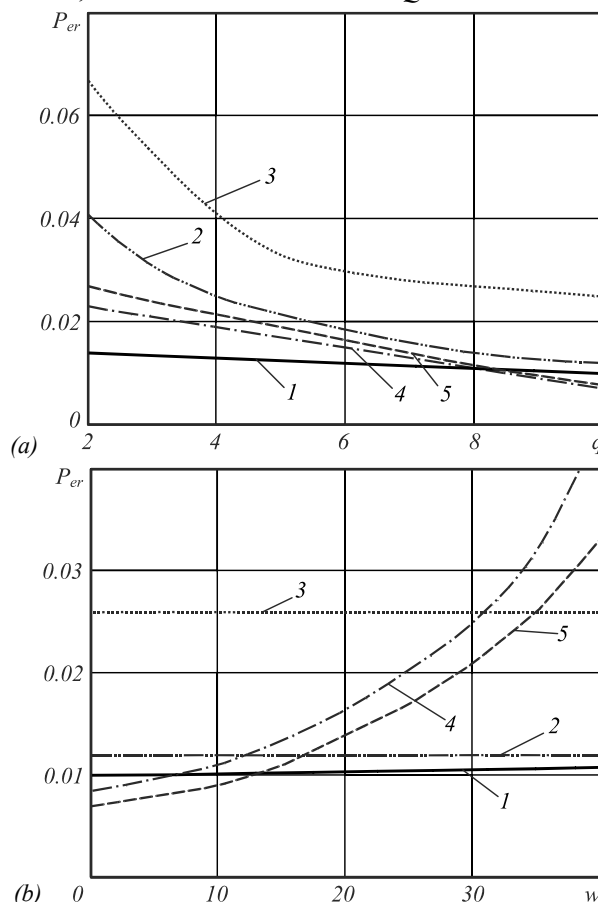


Fig. 2. Dependence of the probability of the false identification of the methods on signal/noise ratio (a) and spatial mismatch of pattern and the object (b)

3. The integral criterion

"computational complexity – recognition quality"

We also compare the methods under study with the integral criterion "computational complexity - recognition quality" proposed in [16]. The numerical value of the criterion is found as the product of the computational cost and the probability of false identification: $\mathfrak{R} = C \cdot P_{er}$, and it characterizes the degree of deviation from the ideal situation: the absence of identification errors and real-time work.

Figure 3a shows the curves of the dependence of the integral criterion on the image sizes at $q=9$, and in figure 3b, on the dimensions of the target object for the same signal-to-noise ratio.

An analysis of the experimental results shows that the best value of the integral criterion is SGI with objective functions MSE and CC. At the same time, the increase in noise has little effect on the behaviour of the integral criterion.

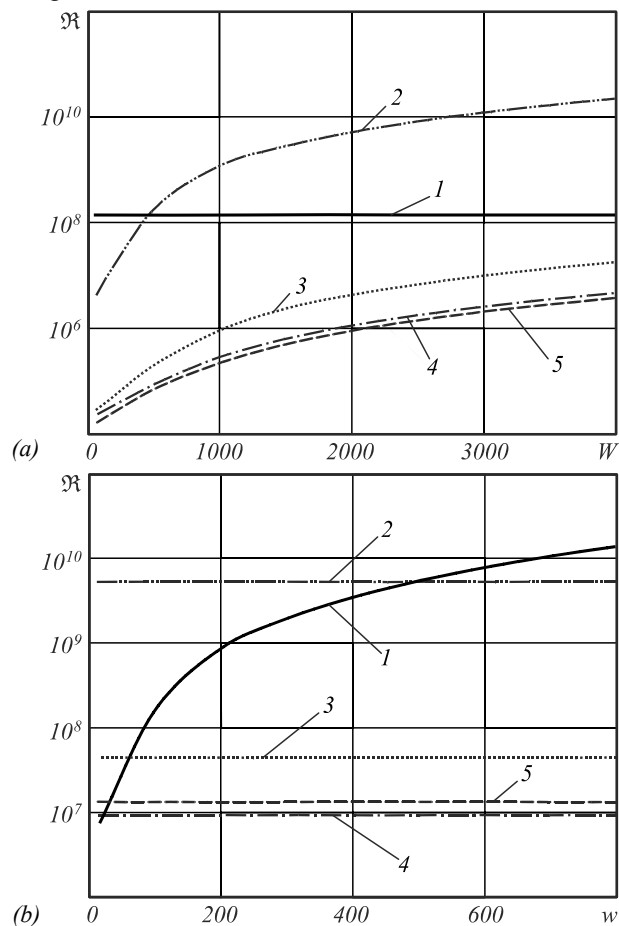


Fig. 3. Dependence of the integral criterion on the image size (a) and object size (b) with signal-to-noise ratio 9

Conclusion

Comparative analysis of the studied identification methods of objects in an image showed that their computational burden depends in different ways on the size of the image. With relatively small image sizes, the CAM ensures a smaller computational complexity, and, when the image sizes are large, the SGI algorithm does. The dependence of the computational complexity of the CEA on the image size is quadratic both in the spatial and frequency regions and approximately two orders of magnitude higher. The computational complexity of the CAM, SGI, and CEA depends weakly on the object size in the frequency region, and for the CEA, it is quadratic in the spatial region. The SGI requires the least computational burden and the CEA the most.

Due to its large sample volume the CEA has the best noise immunity in the spatial region. Here erroneous

identification is specified basically by the interval of change of identification parameters. The SGI also ensures the best noise immunity. However, in this method the probability of correct identification depends on the location mismatch of the sought object and the pattern. The probability of false identification using CAM in noisy conditions is several times higher due to errors in edge detection.

References

- [1] Poltavskii AV, Grinshkun AV. Basics of pattern recognition using computer [In Russian]. *Dvoimie Tehnologii* 2017; 2: 55-66.
- [2] Knyaz VA, Vishnyakov BV, Vizilter YV, Gorbancevich VS, Vigolov OV. Intelligent information processing technologies for navigation and control problems of unmanned aerial vehicles [In Russian]. *Trudi SPIRAN* 2016; 45: 26-44. DOI: 10.15622/sp.45.2.
- [3] Kuznetsov AV, Myasnikov VV. A copy-move detection algorithm based on binary gradient contours. *Computer Optics* 2016; 40(2): 284-293. DOI: 10.18287/2412-6179-2016-40-2-284-293.
- [4] Magdeev RG, Tashlinskii AG. A comparative analysis of the efficiency of the stochastic gradient approach to the identification of objects in binary images. *Pattern Recognition and Image Analysis* 2014; 24(4): 535-541. DOI: 10.1134/S1054661814040130.
- [5] Prett W. *Digital image processing: in 2 volumes*. New York: John Wiley and Sons; 1978.
- [6] Gruzman IS, Kirichuk VS, Kosih VP, Peretyagin GI, Spector AA. *Digital image processing in information systems* [In Russian]. Novosibirsk: NGTU Publisher; 2002.
- [7] Furman YaA, Krevetsky AV, Peredevev AK, Rozhentsov AA, Khafizov RG, Egoshina IL, Leukhin AL. *Introduction to contour analysis and its applications to image and signal processing* [In Russian]. Moscow: "Fizmatlit" Publisher; 2003.
- [8] Tsyppin YaZ. *Information theory of identification* [In Russian]. Moscow: "Fizmatlit" Publisher; 1995.
- [9] Tashlinskii AG. Computational expenditure reduction in pseudo-gradient image parameter estimation. *International Conference on Computational Science* 2003; 2658: 456-462.
- [10] Gonzalez R, Woods R. *Digital image processing*. Upper Saddle River, New Jersey: Prentice Hall; 2012.
- [11] Canny JA. Computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 1986; PAMI-8(6): 679-698.
- [12] Duda RO, Hart PE, Stork DG. *Pattern classification*. New York: Wiley-Interscience; 2001.
- [13] Tashlinskii AG. Pseudogradient estimation of digital images interframe geometrical deformations. In Book: Obinata G, Dutta A, eds. *Vision systems: Segmentation and pattern recognition*. InTech; 2007: 465-494. DOI: 10.5772/4975.
- [14] Tashlinskii AG. The specifics of pseudogradient estimation of geometric deformations in image sequences. *Pattern Recognition and Image Analysis* 2008; 18(4): 700-705. DOI: 10.1134/S1054661808040275.
- [15] Tashlinskii AG. *Estimation of the parameters of spatial deformations of image sequences* [In Russian]. Ulyanovsk: UISTU Publisher; 2000.
- [16] Fadeeva GL. *Optimization of the pseudo-gradient of the objective function in the estimation of inter-frame geometric deformations of images* [In Russian]. The thesis for

the Candidate's degree in Technical Sciences. Ulyanovsk; 2007.

[17] Sebryakov GG, Soshnikov VN, Kikin IS, Ishutin AA. Optimization of parameters of partitioning the analyzed fragment of the image of the scene according to the quality cri-

teria and computational efficiency of recognition of the observed objects [In Russian]. In Book: Technical vision in control systems: materials of scientific and technical conference. Moscow: SAKVOEE Space Research Institute of the Russian Academy of Sciences; 2014: 149-151.

Authors' information

Radik Gilfanovich Magdeev (b. 1987) graduated from Ulyanovsk State Technical University in 2011, majoring in Communication Networks and Switching Systems. Currently he works as the leading Telecommunications Engineer Telekom.ru LLC. Research interests are computer graphics processing, programming and stochastic gradient identification. E-mail: radiktkd2@yandex.ru.

Alexander Grigorevich Tashlinskii (b. 1954) graduated from Ulyanovsk Polytechnic Institute in 1977 (presently, Ulyanovsk State Technical University), majoring in Radio Engineering. Doctor of science, professor, head of Radio Engineering department Ulyanovsk State Technical University. His research interests are currently focused on computer optics, image processing, computer design, and digital photography. E-mail: tag@ulstu.ru.

Received August 10, 2018. The final version – March 22, 2019.
