Scientific
Research
Publishing

The Use of Fourier Descriptors for the Classification and Analysis of Peripheral Blood Smears Image

Rimma Tomakova, Vitaliy Komkov, Evgeniy Emelianov, Maksim Tomakov

South West State University, Kursk, Russia
Email: rtomakova@mail.ru, vitaliykomkov@gmail.com, david46@ramble.rum, Tomakovmv@rambler.ru

Abstract

The article discusses the use of Fourier descriptors for the analysis and classification of blood cells. A model describing the contour boundaries in the form of two-dimensional numerical sequence Fourier descriptors. The influence of the shape and orientation of the figures on the parameters of the Fourier descriptors. Explore ways to ensure the invariance of the Fourier descriptors with respect to geometric transformations. A model of the graphical representation of the Fourier descriptors of computer graphics tools. A method of forming a space of informative features based on Fourier descriptors for the neural network, classifying the contours of borders image segments.

Keywords

Fourier Descriptors, Image Processing, Analysis of the Spectrum Boundaries, Space of Informative Signs, Recognition, Classification of Objects

1. Introduction

Currently, in artificial intelligence systems, the segmentation and classification of objects on medical images are widely used Fourier descriptors.

Analysis of existing approaches has shown that the known methods for determining the image descriptors and their segments such as: SIFT (Scale Invariant Feature Transform), SURF (Speeded Up Robust Features), PCA-SIFT (PCA-Principal Component Analysis) etc. intended for use in intelligent systems and similar search similar images, but not in the classification systems. The implementation of these algorithms requires a large volume of samples [1]. However, with regard to medical images provide necessary training sample is not always possible. These restrictions are associated with the presence of disease and

the complexity of the experiments on living objects [2] [3].

2. Implementation of the Method

In practice, the detection of blood cells are faced with a major obstacle, typical of medical applications of computer vision—a large variability in the images [4].

To carry out the analysis and classification of blood cells, shown in the **Figure 1**, the image must be analyzed and segmented and binarized [5] [6] [7]. The task is to detect all the blood cells in the image and representation of pixels "units", and pixels other than the blood corpuscles, "zeros".

For the analysis of black and white segments used Fourier descriptors. When identifying and segmenting images used spectral analysis of two-dimensional numerical sequence obtained by sampling the curve $L(x, y)$ describing the image segment of the border. The boundary traversal starting from an arbitrary pixel (x_0, y_0) belonging curve defines this sequence.

Performing bypass loop counterclockwise, obtain a set of coordinates of the border points $\{(x_0, y_0), (x_1, y_1), \cdots, (x_{k-1}, y_{k-1})\}$ where K —the number of pixels on the boundary. These coordinates can be represented in parametric form:

$$x_k = f_1(k) \tag{1}$$

$$y_k = f_2(k) \tag{2}$$

In this case, the boundary can be represented as a series of
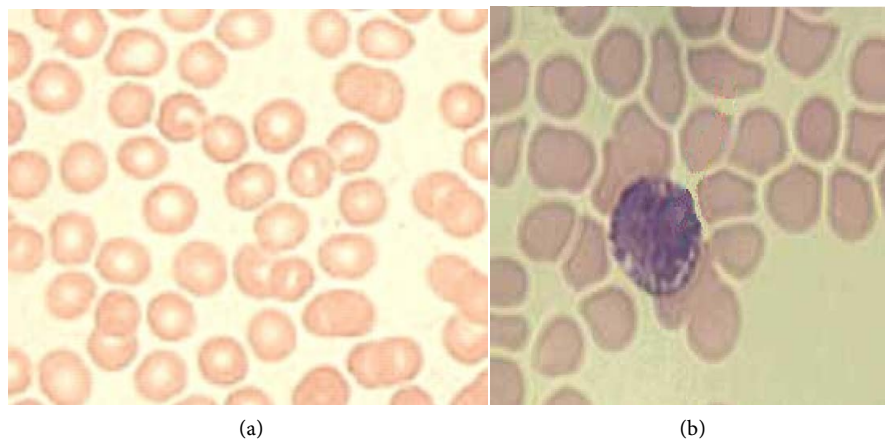
$$s(k) = \left[f_1(k), f_2(k)\right], \tag{3}$$

where the number of k pixel is $k = \overline{0, K-1}$.

Thus, we should regard each pair of coordinates as a complex number of the form:

$$\dot{z}(k) = f_1(k) + j f_2(k), \tag{4}$$

for $k = \overline{0, K-1}$.

Thus, consider x and y as real and imaginary parts of complex numbers sequence. This representation has an advantage: it allows you to reduce the two-

Figure 1. Blood cells. (a) Erythrocytes; (b) Smear with verdure.

dimensional problem to a one-dimensional spectral analysis [8] [9].

Discrete Fourier transform of a finite sequence $\dot{z}(k)$ is given by:

$$z(k) = \frac{1}{K}\sum_{u=0}^{K-1} a(u) e^{j2\pi uk/K}\,, \tag{6}$$

for $k = 0,1,2,\cdots,K-1$.

To investigate the spectral maps used architectural software package, developed in the environment of Matlab 7.10. This software package allows you to define and visualize the real, the imaginary and the amplitude component of the contour of the spectrum, to analyze the dependence of the change in the shape of segments of the Fourier descriptors used [10]. Analysis of the spectrum segments boundaries conducted on the basis of (4) showed that the spectrum Fourier descriptors has a symmetric structure (**Figure 2**).

This means that the Fourier descriptors forms pairs by frequency and each descriptor on the negative real axis corresponds to the handle on the positive half with exactly the same frequency coordinates. The amplitudes of these descriptors may be in various ratios.

Note that to restore the boundaries of contours are not used all but only the first of *M* Fourier coefficients. As a result of the recovery we have approximation $\hat{z}(k)$:

$$z(k) = \frac{1}{K}\sum_{u=0}^{K-1} a(u) e^{j2\pi uk/K}\,, \tag{6}$$

for $k = 0,1,2,\cdots,K-1$.

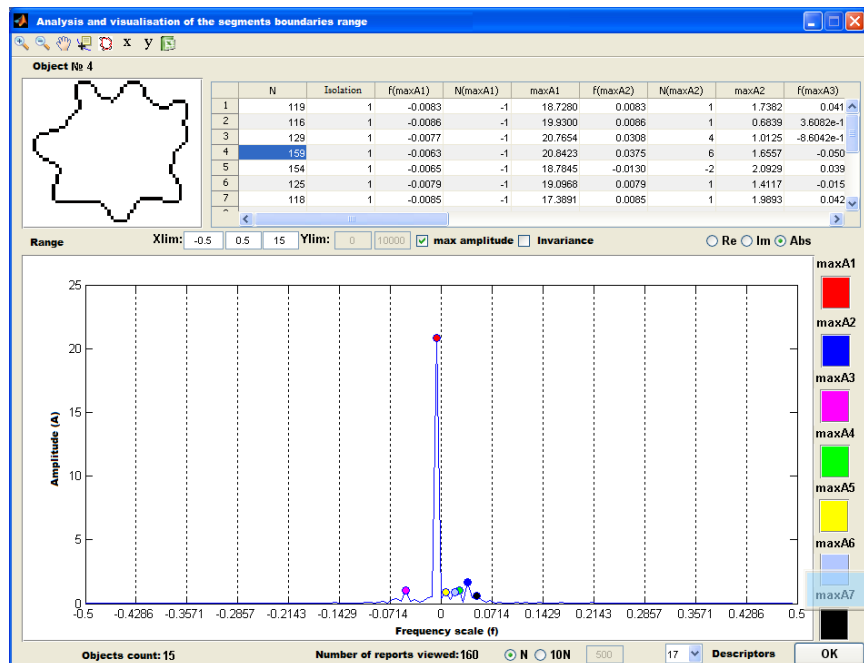Analysis of the research shows that high-frequency components describe the

Figure 2. Interface window in the calculation of the Fourier descriptors shows the amplitude spectrum.

fine details, while the low-frequency components determine the overall shape of the border. Therefore, less than *M*, the more detail is lost the boundary [11] [12] [13].

Figure 3 shows the boundary of an arbitrary contour, consisting of $K = 130$ points, and also shows the stages of its recovery with the help of Equation (7) for different values of *M*. The results obtained with the help of a software package, developed in the environment of Matlab 7.10.

Starting from about 33 *M* values have reduced border corner points are beginning to appear. When $M = 98$ is restored almost an exact copy of the original. This implies that several lower-order coefficients are sufficient to describe the general shape of the boundary, but for accurate recovery of sharp parts, for example, corners and straight segments, requires a much larger number of high-order terms [9].

Analysis of the results shows that the Fourier descriptors synthesize space informative features in the area of spatial frequencies to identify the contour shape.

Upon receipt of the boundaries of the spectrum based on (5) is necessary to achieve invariance descriptors to shift and turn the segment, as well as the possibility of selecting segments with different scales.

In the analysis of medical images often arises the problem of segmentation of a certain class. Often these segments derives from a parent segment, the result obtained by converting the original scale. Zooming can be both informative source, and be a hindrance.

3. Invariance with Respect to Scale

If the contour can be scaled by the factor *α*, then all Fourier descriptors are also can be scaled by *α*. If the contour of the segment has non-zero area and can be traced counterclockwise, the first factor (first harmonic) is always non-zero.
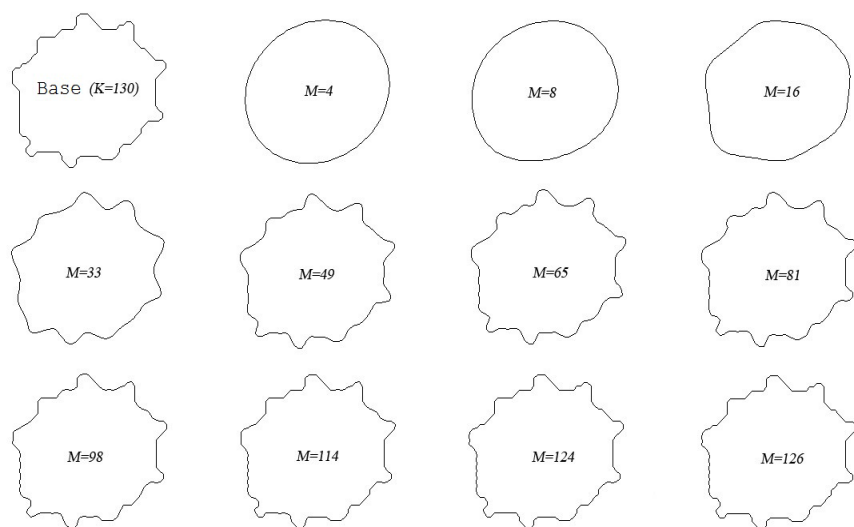
Figure 3. Stages of recovery circuit boundaries based on Fourier descriptors: *M*—the number of Fourier coefficients used in the restoration.

Thus, scaling the Fourier descriptors is achieved by normalization of the absolute relative Fourier component of the first descriptor.

Providing invariance under scale is relevant for comparison and classification of segments in various photographs taken with different scales

4. Invariance with Respect to Shift

A single factor a_0 (zeroth harmonic) limits the position of the object (contour). All other factors are invariant with respect to displacement. Invariance under transformations "shift" can be achieved by equating to zero the zeroth coefficient (the first coefficient, which has zero or one index depending on the software package; in the package of spectral transformation Matlab 7.10 it is one).

Invariant with respect to rotation. Invariant with respect to rotation defined by absolute value of the Fourier descriptors. This means that the representation of descriptors in the form of complex vector modules is invariant with respect to rotation. Therefore, consider the problems of classification, only peak values of the spectrum should be used.

The optimal choice of the Fourier descriptors used in the analysis of the boundaries of image segments

It is necessary that the frequency range corresponding to a specific Fourier descriptor number is independent of the number of counts in the loop whose spectrum is calculated, to have adequate recognition system. In general, the various circuits contained different numbers of samples. At the same sample rate and different sample quantities in the contours, this requirement is violated.

Based on the research of Fourier descriptors it developed a method of forming a space of informative signs. In the first step specifies the total number of samples in the circuits, which must be the same for all the control loops and training samples. This number is determined based on the results of statistical studies. For microscopic photographs of the vessels, the number of samples (pixels) in the contours varies from 500 to 3000. In order to be able to compare the discrete frequency samples corresponding to different circuits the contours, that have same number of samples, are needed. To equalize the number of samples in the analyzed circuits sample is necessary to make the number of samples in each circuit to a maximum. This is accomplished by supplementing the high-frequency part of the spectral band zeros. For example, suppose the number of samples in spectrum of each segment boundary contour maximum, equal to 3000. In this case in the space of signals will appear virtual counts between real counts.

Thus, we have the following chain of transformations: Kmax → spectrum→ zero padding spectral readings at high frequencies up to Kmax.

Note that the Fourier descriptors feature is that the amplitude associated with frequency. Therefore, any changes in the frequency of the signal lead to real changes in the amplitude of spectral components. The criterion of adequacy of any changes in the frequency domain is inverse Fourier transform, and the re-

spective differences between the direct and inverse Fourier transform.

Therefore, in order to save energy equivalent between the space of signals and the space of frequencies, it is necessary to reduce the amplitude of descriptors in proportion to the number of intermediate samples appearing in the signal space. Thus, on the second phase all spectral components in the spectrum of the i-th circuit must be multiplied by Kmax/Ki value where Kmax = 3000, Ki—the number of samples in the i-th circuit.

Figure 4 shows the spectrum circuits before (a) and after modification in the spectral region (b).
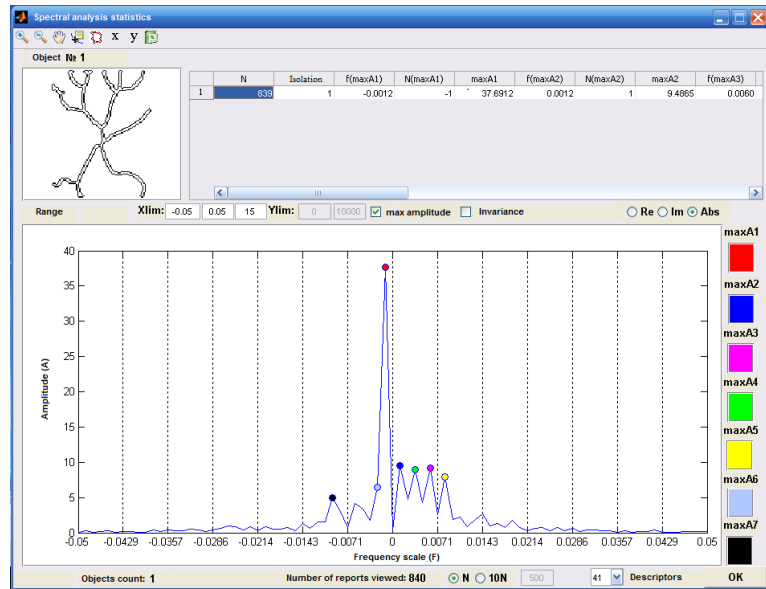
(a)
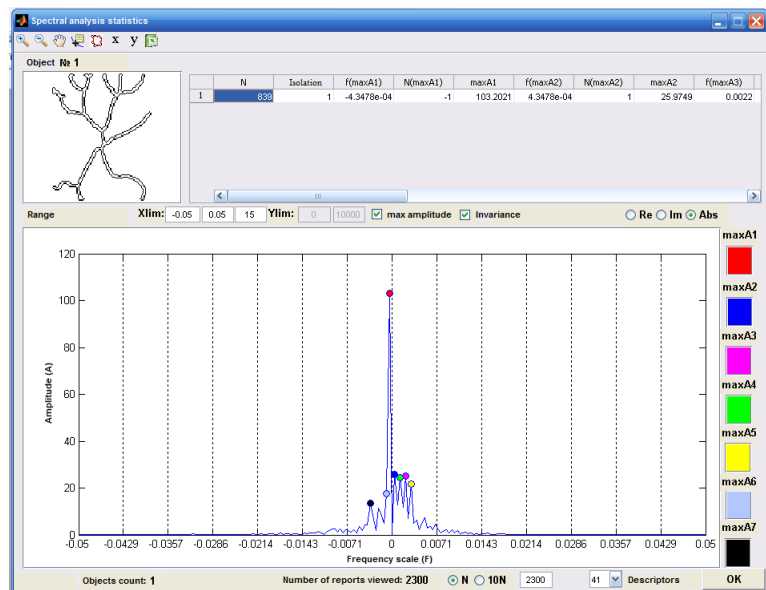
(b)

Figure 4. Spectrum of contour segment boundaries before modification (a) and after modification (b).

The third stage-the optimization of the number of descriptors used. For the optimal choice of the analyzed descriptors (the optimum is meant to minimize the number) necessary to implement the inverse Fourier transform of the modified spectrum circuit and compare it with the original contour.

Considering that, the Fourier transformation is reversible, and using Equation (7), Fourier descriptors can restore the circuit boundaries of the test segment. In this case for recovery of the circuit use same amount of descriptors as number of counts that were obtained on the contour. If part of the descriptors set equal to zero, then, using a neural network classification model, is lead to a considerable simplification of decision-making model.

To evaluate the efficiency of this method is necessary to assess the information loss arising from equating to zero part of the descriptors. For this purpose, it is proposed to compare the original contour and contour restored on a limited set of descriptors. After the transition from K descriptors to M descriptors ($M < K$), we obtain the parametric curves that reflect the geometry of the segment boundaries in the form of

$$\tilde{x}_k = \tilde{f}_1(k) \tag{8}$$

and

$$\tilde{y}_k = \tilde{f}_2(k) \tag{9}$$

Examples of the starting and reconstituted by M descriptors of parametric curves (1), (2) and (8), (9) are shown in **Figure 5**.

Since the scale of the curves are (**Figure 5**), the area of S circuit defined as

$$S = \frac{1}{K}\sum_{k=0}^{K-1}\left|y(k(\xi,\zeta)) - y(k(\xi,\zeta))\right| \to \min, \tag{10}$$

where K—the number of samples in the sample circuits, ξ and ζ—the value (reconstituted circuit), located at a minimum distance from the k (source circuit), and for bitmaps $k = k^*$.

Figure 6 shows plots of the information loss dependence of the number of descriptors used when restoring circuit segment boundaries.

Optimizing the number of descriptors performed based on Kettel's criterion
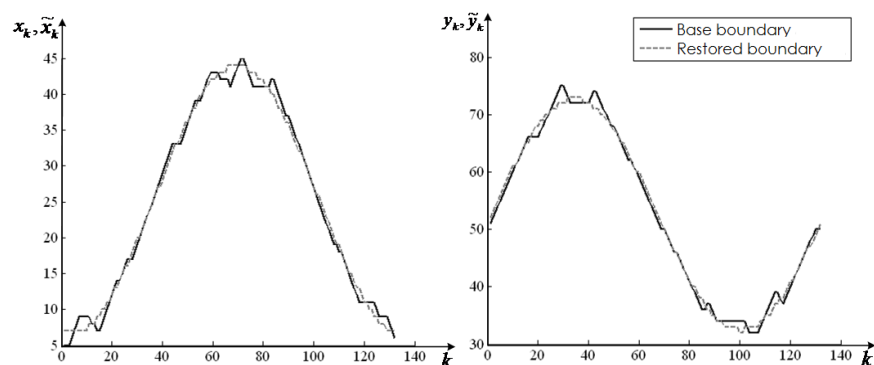
Figure 5. Parametric curves of the original and the reconstructed contour border.
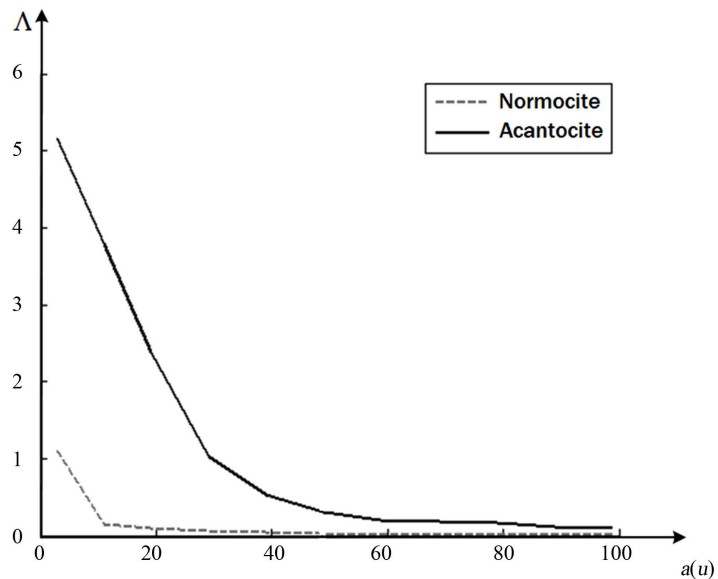
Figure 6. Plots of the information loss dependence of the number of descriptors used for the two circuits.

[14]. According to this criterion it is defined a place on the charts, where the decrease of the criterion Λ left to right is slowing as much as possible. In accordance with this criterion, appropriate to use 40 descriptors for the test circuit.

Using the developed methods and algorithms for spectral analysis of the boundary curves, we obtain an adequate description of the curve of any type end, a priori a certain length of Fourier descriptors sequence.

5. Conclusions

A method for forming a space for informative features neural network classifies freeform curve, based on the definition of the Fourier descriptors of the corresponding, circuit was developed.

A model describing the contour boundaries in the form of two-dimensional numerical sequence Fourier descriptors was proposed.

Abstract models of various geometric shapes based on Fourier descriptors were examined; the influence of the shape and orientation of the figures on the parameters of the Fourier descriptors was found; it offered a model of the graphical representation of the Fourier descriptors, allowing forming a space of informative features for intelligent image classification systems.

We explored ways to ensure the invariance of the Fourier descriptors with respect to geometric transformations.

On typical medical images, containing from 500 to 3000 pixels, the proposed model should reduce the space of informative signs in order to further use of the neural network pattern recognition.

A method for forming a space of informative features for neural network, classifying the contours of borders image segments, has been developed. The method is based on determining the Fourier descriptors for the contour, corres-

ponding to the boundary segment, allowing obtaining an adequate classification model, regardless of the number of samples in the circuits.

References

[1] Sonka, M., Hlavac, V. and Boyle, R. (2014) Image Processing, Analysis, and Machine Vision. *Cengage Learning*, 875.

[2] Schwartz, W. (1987) Artificial Intelligence in Medicine: Where Do We Stand. In: Schwartz, W., Patil, R. and Srolovits, P., Eds., *New England Journal of Medicine*, **316**, 685-688.

[3] Semmlow, J.L. and Griffel, B. (2014) Biosignal and Medical Image Processing. CRC Press, 592.

[4] Babenko, A., *et al.* (2014) Neural Codes for Image Retrieval. *European Conference on Computer Vision*, Springer International Publishing, 584-599.

[5] Kozinets, G.I. (1997) The Study of the Blood System in Clinical Practice. In: Kozinets, G.I. and Makarov, V.A., Eds., *Triad-X*, 480.

[6] Tomakova, R.A. (2010) Analysis of Histological Images by Morphological Operators, Synthesized Based on Fourier Transform and Neural Network Modeling. In: Tomakova, R.A., Filist, S.A., Gorbatenko, S.A. and Shvetsova, N.A., Eds., *Biotechnosphere*, **3**, 54-60.

[7] Sharon, E. (2000) Fast Multiscale Image Segmentation. In: Sharon, E., Brandt, A. and Basriy, R., Eds., *Proceedings IEEE Conference on Computer Vision and Pattern Recognition*, 70-77.

[8] Candemir, S., *et al.* (2014) Lung Segmentation in Chest Radiographs Using Anatomical Atlases with Nonrigid Registration. *IEEE Transactions on Medical Imaging*, **33**, 577-590.

[9] Reed, T.R. (2005) Digital Image Sequence Processing, Compression, and Analysis. In: Reed, T.R., Ed., CRC Press, Boca Raton, 272 p.

[10] Kubrin, S.S. (2005) Comparison of the Method of Geometric Moments and the Method of Fourier Descriptors in the Text Recognition Tasks. In: Kubrin, S.S., Fedchenko, N.A. and Isaev, A.B., Eds., *Information-Analytical Bulletin of the Moscow State*, Mining University, 106-108.

[11] Deacons, V.P. (2006) MATLAB 6.5 SP1/7/7 SP1/7 SP2 + Simulink 5/6. Tools of Artificial Intelligence and Bioinformatics. In: Deacons, V.P. and Kruglov, V.V., Eds., Solon-Press.

[12] Borisov, S.A. (2011) The Certificate of Registration of the Computer Program No. 2011613254. Software Morphological Spectral Image Analysis. In: Borisov, S.A. and Tomakova, R.A., Eds., *Rospatent*; The Applicant and the Right Holder SWSU, Number 2011611389.

[13] Khalafyan, A.A. (2007) STATISTICA 6. Statistical Analysis. In: Khalafyan, A.A., Ed., OOO Binom Press, 512.

[14] Tomakova, R.A. (2012) The Method of Processing and Analysis Complex Images Based on Built-In Functions MATLAB Environment. In: Tomakova, R.A. and Filist, S.A., Eds., Herald of Chita State University, 3-9.