Smoothing Algorithm for Magnetometric Data

Mariya PASHKOVA, Sofya ILINYKH,
Saint-Petersburg State University
St. Petersburg, Russia
Pashkova.Mariya@rambler.ru, Ilinykh-Sofya@mail.ru

Natalia GRAFEEVA
Saint-Petersburg State University
St. Petersburg, Russia
N.Grafeeva@spbu.ru

***Abstract*— This article presents the problem of searching useful magnetic anomalies by using the magnetometric survey methods. Metal debris (spot anomalies) often create local interference that makes it difficult to find vast spatial anomalies caused by the remnants of ancient buildings, such as walls of houses, wells, dugouts, etc. The main purpose of this study is to develop an algorithm that eliminates such interference. We explore the methods of working with magnetometric data and design an algorithm which is based on seeking spot anomalies and using the arithmetic method of smoothing spatial data. We test the final algorithm on many real datasets, obtained during excavations in the Crimea, and it shows good results.**

Keywords— magnetometric, spatial data smoothing, arithmetic mean of spatial data

I. Introduction

In archaeology there are different methods of determining the location of future excavation sites. One of those methods, *magnetometric survey*, is based on magnetic properties of hidden underground objects. The main advantage of this method is a relatively low cost.

The process of magnetometric survey is described in [6]. Magnetometric survey is usually performed on a square surface area of proposed excavation site with the size ranging from 50 by 50 meters to 100 by 100 meters**.** That square area is called a *quadrangle*. At first, a modulus of magnetic induction of the geomagnetic field is measured at previously determined observation points. Then, collected data is preprocessed (e.g. to take into account global geomagnetic field changes). Then, researchers build a map of a quadrangle, estimate the probability of finding objects of archaeological value. Finally, a conclusion on conducting excavations is made. Sometimes, anomalies (deviation from the average value of magnetic field), created by past human activity, can be spotted on a map.

Often, small metal junk creates interferences represented in numerical data by sharp value drops causing difficulties in searching for vast spatial anomalies created by ancient buildings (houses, wells and burial mounds). The main objective of this study is developing and testing an algorithm which will allow smoothing out sharp data drops thus eliminating interferences caused by metal junk.

II. Data Description

Provided data is a result of a magnetometric survey that was conducted on excavation sites in Crimea [3]. Magnetometric data is an example of spatial data: each data set consists of 15000 three-dimensional points. The first two components are coordinates; the third one is the difference between value of magnetic field at a given point and the average value of magnetic field at the quadrangle rounded to the nearest whole (measured in nT). The point coordinates are measured with 0.5 meters increments. Such spatial data can be represented as a two-dimensional image where the third component plays the role of color.

The visualization of this data can be created in RStudio IDE which is also capable of providing the data statistics useful for some algorithms implementation. Table I-II and Fig.1 provide examples of input data, visualization and statistics calculation in RStudio IDE.

TABLE I. An Example of Initial Data

X	0	0	0	0	0	0	0	0	…
Y	49.5	50	50.5	51	51.5	52	52.5	53	…
value	-3	3	156	-3	4	0	-6	-11	…

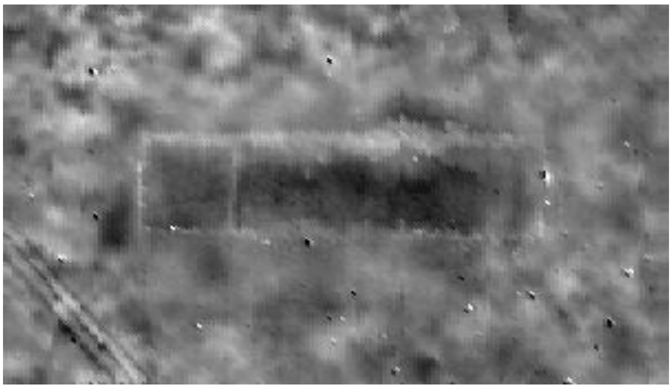

Fig. 1. Dataset visualization.

TABLE II. DATA STATISTICS

	x	y	Value
Min	0.0	49.0	-322.000
1st Qu.	12.5	62.0	-6.000
Median	25.0	75.0	-1.000
Mean.	25.0	75.0	-1.754
3rd Qu.	37.5	88.0	2.000
Max	50.0	101.5	285.000

III. OVERVIEW OF METHODS FOR WORKING WITH MAGNITOMETRIC DATA

There are three groups of methods are usually used to process magnetometric data.

A. Working with data as an image

Magnetometric data can be easily visualized [3]. For example, magnetic value can serve as a color of the black-to-white gradient. This way researchers can process data as an image, apply filtration [10], stretching, etc. Some of these methods are presented in [3].

There are two different ways of processing images within the context of the task:

- Methods based on processing pixels directly. For example, logarithmic and power transformations, application of Gaussian filter [8] and Sobel operator [9], etc.
- Methods, based on modification of the Fourier spectrum of the image [9]

B. Working with numerical data, based on its nature

Considering the nature of data, it can be processed as numbers. In each point x of quadrangle input data is represented as a sum

$$B(x) = R(x) + P(x) + A(x)$$

where R is a level of regional background magnetic field, P – interference, A – local anomalies.

Applying certain algorithms (analytical continuation of the field into the upper half plane, Kalman filter, etc.) the researcher is trying to distinguish and deduct the background and the interference to find important local anomalies. The main source of information about those methods is [2].

C. Other approaches

Lately, a variety of methods (clusterization [7], patterns recognition [11], etc.), that can be applied to magnetometric data, but do not take in consideration the nature of the data, have been suggested.

Often image processing methods cannot provide the desired results. Application of those methods sometimes causes strong distortion of the image (reasons include aforementioned metal objects), which hinders the search for useful spatial anomalies. Methods, based on the nature of the data, are interesting and sometimes useful, but are costly. At the same time simple disposal of interference, created by metal objects, can drastically change the visual image of the analyzed data. With that in mind in this study attention was focused on finding an algorithm, which processes data to eliminate interference that hinders the search of vast space anomalies.

IV. CHOOSING A METHOD FOR SMOOTHING MAGNETOMETRIC DATA

In Fig. 2 black dots surrounded by white areas are clearly visible (in numerical data they are represented by large deviations from average value of the quadrangle). These are the point anomalies, caused by metal junk or instrumentation errors. Usually, they cover no more than 1-2% of the whole area. They are not always objects of archaeological research, but they can obstruct analysis of more significant objects.

So, the idea is to smooth the values in such points. Then, small meaningful deviations will be more noticeable when visualized.

This can be achieved, for example, by a simple method of arithmetic mean, but instead of time series, space surrounding the interference points can be used.

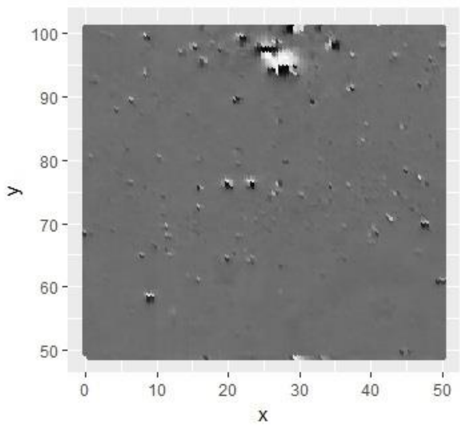

Fig. 2. Example of input data with point anomalies

V. DEVELOPMENT OF ALGORITHM

During data analysis it was observed that in places of metal junk concentration of large deviations occur, both positive and negative. At the same time, points where such situation happens make up a small portion of all points in the quadrangle. There are two approaches for finding point anomalies based on these ideas.

A. Frequency-based approach

First approach is based on sorting magnetic value by frequency of appearance on a set area. Afterwards points, that appear rarely enough (i.e. make up no more than M percent of the area), are selected. Table II shows a table of magnetic values and their corresponding frequencies with values, selected to be smoothed, highlighted. In this example M equals 2%. Fig. 3 shows values of the examined area with selected values marked in dark grey.

B. Deviation-based approach

In the second approach points are considered to be interference when their value deviates from the average by more than parameter δ.

Fig. 4 shows selected points with δ = 9. The average value is -3.5.

After the interference points have been found, their values should be smoothed. Following solution has been considered: points are assigned an average value of some neighborhood that does not include point anomalies.

Firstly, two extreme cases have been implemented. In the first case an interference point was assigned a value of the closest (using the Euclidean distance) point that is not a subject to smoothing. In the second case, the average value of the quadrangle is assigned to all interference points. As can be seen in Fig. 5-7, results of applying these extreme cases are not what we are looking for: smoothing is too rough and inaccurate, distortions are still visible and local anomalies become blurred, thus it does not solve our problem. In both cases the interference points were found using deviation method.

TABLE III. MAGNETIC VALUES AND A PERCENT OF AREA COVERED BY EACH VALUE

Magnetic values	-14	-13	-9	3	4	-16	-12	-10	-6	-3	5	6	-8	2	7	-1	-4	-5	0	-2
Percent of area	2%	2%	2%	2%	2%	3%	3%	3%	3%	3%	3%	5%	5%	5%	6%	6%	8%	9%	13%	16%

2	2	5	6	6	0	-7	-5
-3	-1	0	5	4	2	-1	-2
-5	-2	-2	3	-2	0	0	0
-2	-5	-4	0	-8	-4	-1	-3
-2	-6	-7	-2	-10	-10	-4	-6
-5	-5	-12	-5	-8	-16	-9	-4
-1	-2	-7	-4	-8	-14	-13	0
1	0	-2	-2	-5	-12	-16	-7

Fig. 3. Example of finding interference point using a frequency method, parameter M = 2%

2	2	5	6	6	0	-7	-5
-3	-1	0	5	4	2	-1	-2
-5	-2	-2	3	-2	0	0	0
-2	-5	-4	0	-8	-4	-1	-3
-2	-6	-7	-2	-10	-10	-4	-6
-5	-5	-12	-5	-8	-16	-9	-4
-1	-2	-7	-4	-8	-14	-13	0
1	0	-2	-2	-5	-12	-16	-7

Fig. 4. Example of finding interference point using a deviation method, parameter δ = 9

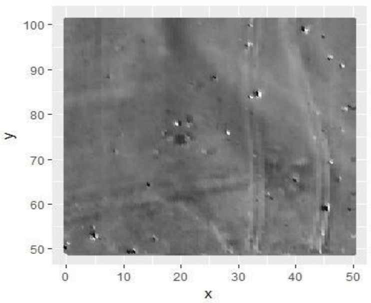

Fig. 5. Input data

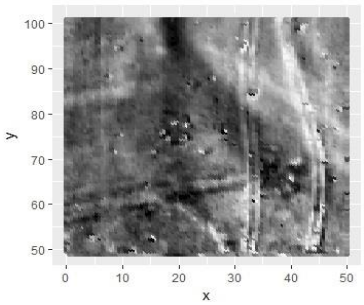

Fig. 6. Single-point smoothing result

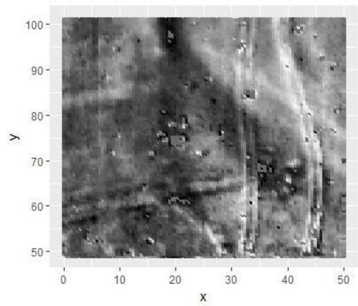

Fig. 7. Smoothing using all quadrangle

2	2	5	6	6	0	-7	-5	-5	-5	-6
-3	-1	0	5	4	2	-1	-2	-5	-3	-4
-5	-2	-2	3	-2	0	0	0	-2	-1	-3
-2	-5	-4	0	-8	-4	-1	-3	-3	-1	2
-2	-6	-7	-2	-10	-10	-4	-6	-3	-3	3
-5	-5	-12	-5	-8	-16	-9	-4	1	-1	2
-1	-2	-7	-4	-8	-14	-13	0	1	-2	2
1	0	-2	-2	-5	-12	-16	-7	-5	-3	-1
1	-1	-2	-2	-3	-7	-16	-12	-11	-12	-9
-1	-2	-5	1	-3	-4	-13	-18	-11	-11	-6
-1	-3	-1	0	1	-2	-9	-16	-12	-7	-7
-1	-3	-2	2	3	0	-4	-11	-12	-4	3
-2	-2	2	1	-1	-2	2	-2	-9	-9	2
-2	-2	0	3	3	0	0	-1	-12	-13	1

Fig. 8. Example of located interference points and their neighbourhoods.

During the work it was found that in most cases 8 is an optimal number of points in a neighborhood. Fig. 8 shows interference points (marked in gray) and their corresponding neighborhoods (marked in light gray). The interference points were found using the deviation method, where parameter M is equal 9.

Fig. 9-10 present results of the resulting algorithm with different values of parameters δ and M. It can be concluded that frequency method deforms the initial image less, but it is less effective in getting rid of interference points

For the next step of the work it was decided to divide the initial quadrangle into smaller parts and then apply aforementioned approaches to each of them. Figures 11-12 show the results of dividing the quadrangle into 25 parts.

Results are improved and the frequency method performs better than the deviation method. Thus, by dividing initial data into separate parts and smoothing rarely occurring values, desired results can be achieved.

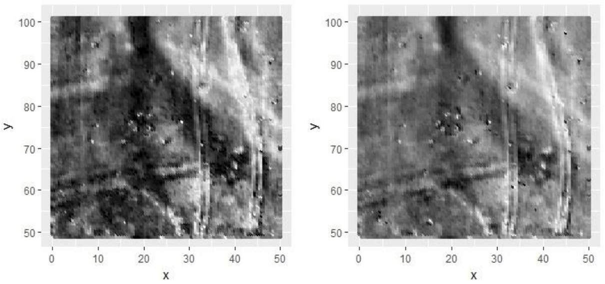

Fig. 9. Frequency method, M = 0.45% and M = 4%

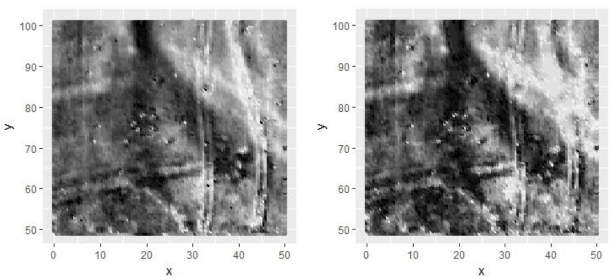

Fig. 10. Deviation method, δ = 13 and δ = 9

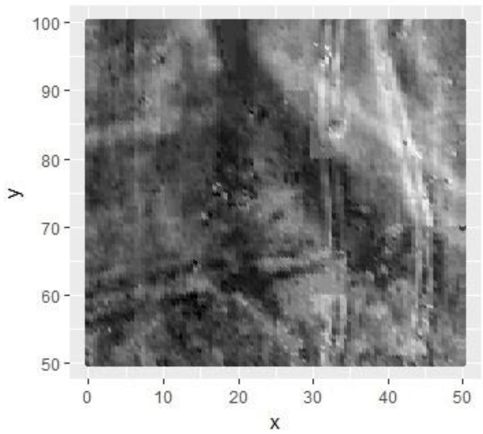

Fig. 11. Deviation method, δ = 8

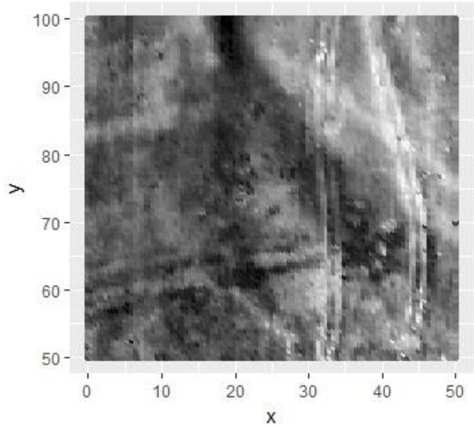

Fig. 12. Frequency method, M = 7%

VI. Conclusion

During the work spatial magnetometric data was smoothed using arithmetic average values helping in identifying local magnetic anomalies.

As can be seen in Fig. 13-14, developed algorithm processes initial data in such way, that an archaeology specialist would have no trouble determining the contours of the object and deciding if the excavation is advisable at the sector in question.

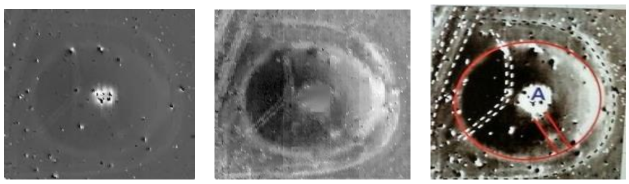

Fig. 13. Initial data, the result of the algorithm and interpretation of archeologists[1]

An advantage of the algorithm is that it transforms numbers, not images. Therefore it can be used not only by itself, but also as preprosessing step for other methods, for example, clustering[4],[5] or filtering[10].

Sample data Result of Algorithm

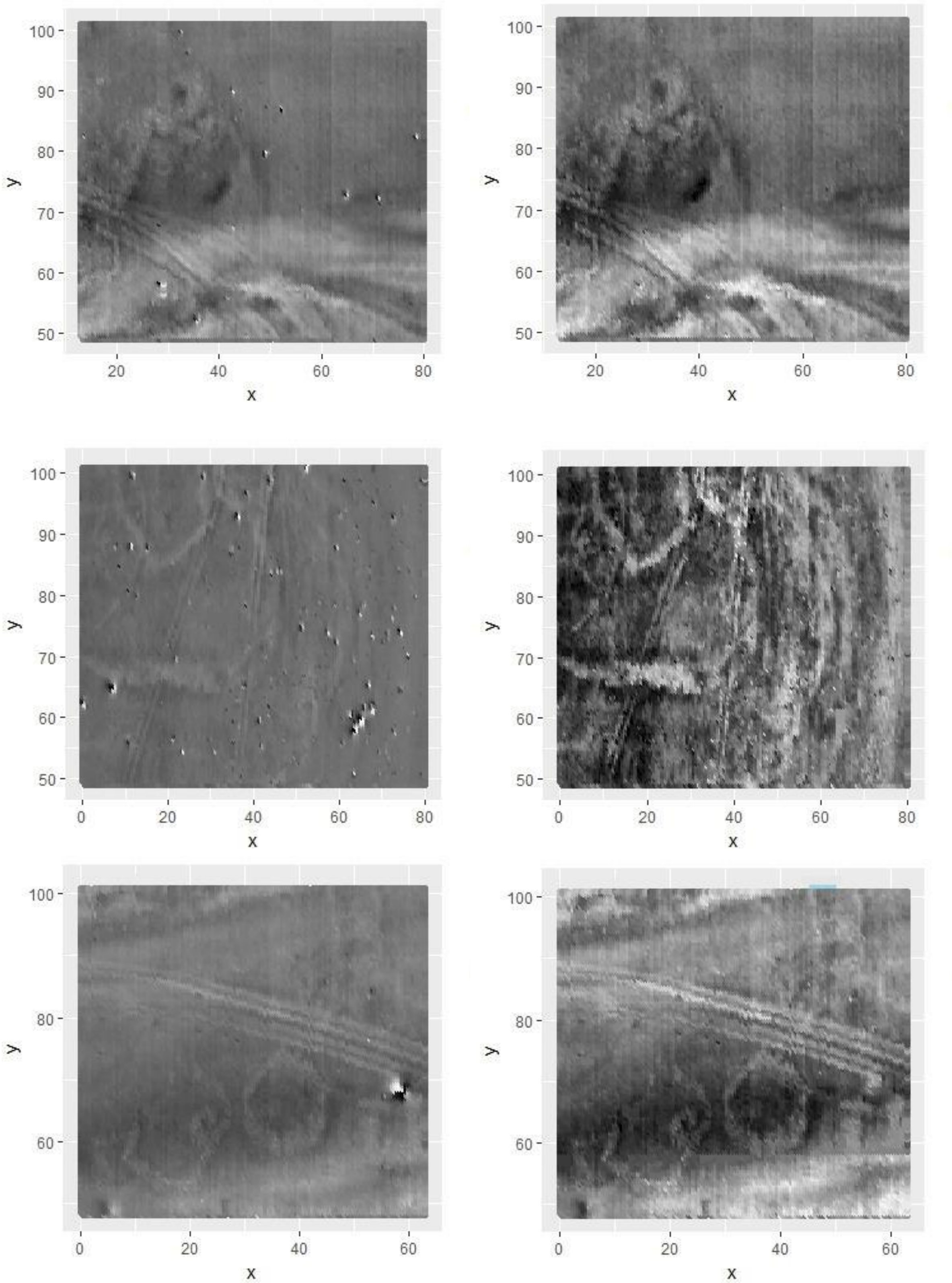

Fig. 14. Another results of testing algorithm on real data

ACKNOWLEDGMENT

The authors would like to thank the Low Field Magnetic Resonance Laboratory in Saint Petersburg State University for provided data, which were obtained as a result of excavations in the Crimea.

REFERENCES

[1] V.A.Kutajsov, T.N.Smekajlov, Materials for archaeological map of Crimea: History and archaeology of Northwestern Crimea, Simferopol: Phoenix Company, 2014. [in Russian]

[2] Dudkin V.P. and Koshelev I.N., Method for complex interpreting the results of magnetometric survey of archaeological landmarks, Vostochnoevropeisk. Arkheol. Zh., No. 3(6), 2002.

[3] V. S. Mikhailova, N. G. Grafeeva, E. G. Mikhailova, A. V. Chudin, Magnetometry Data Processing to Detect Archaeological Sites, Pattern Recognition and Image Analysis, 2016, Vol. 26, No. 4, pp. 789–799. ©Pleiades Publishing, Ltd., 2016.

[4] S. V. Belim, P. V. Kutlunin, Boundary Extraction in Images Using A Clustering Algorithm, Computer Optic, Vol. 39, No. 1, pp. 119–124, 2015. [in Russian]

[5] Elena Volzhina ; Andrei Chudin ; Boris Novikov ; Natalia Grafeeva ; Elena Mikhailova, Discovering geo-magnetic anomalies: a clustering-based approach, 2016, Page(s):1–7 Intelligence, Social Media and Web (ISMW FRUCT), 2016 International FRUCT Conference

[6] V. K. Khmelevskoi, Geophysical Methods for Investigating the Earth Crust, Dubna: Dubna Univ., 1999. [in Russian]

[7] Barsegjan A.A., Kuprijanov M.S., Holod I.I., Tess M.D. and Elizarov S.I., Data and process analysis, St. Petersburg: BHV-Petersburg, 2009. [in Russian]

[8] Duda and P. Hart, Pattern Classification and Scene Analysis, New York: John Wiley and Sons, 1973.

[9] Jian-Lei Liu Da-Zheng Feng, “Two-dimensional multi-pixel anisotropic Gaussian filter for edge-line segment”, Image and Vision Computing archive, vol. 32, Jan. 2014, pp. 37–53.

[10] R. Gonzalez and R. Woods, Digital Image Processing, New Jersey: Prentice Hall, 2008.

[11] Mesteckiy L. M., Mathematical methods of pattern recognition. Lecture course, Moscow: Moscow State University, Faculty of Mathematic and Cybernetics, 2002. [in Russian]