Automation Methods for Processing Medical Images Based on the Application of Grids

Liudmyla Vasylieva[1[0000-0002-9277-1560]], Oleksandr Tarasov[2[0000-0002-0493-1529]]

Donbass State Engineering Academy, Akademichna str.,72, Kramatorsk, 84313, Ukraine
[1]vasileva.dgma@gmail.com
[2]alexandrtar50@gmail.com

Abstract. The problem of reducing the complexity of the process of automated recognition and obtaining a quantitative assessment of the graphic objects parameters in medical images are considered. A method for applying a grid to an image is proposed. It makes possible to provide acceptable accuracy of calculations, a high level of generalization of data, as well as reduce the complexity of calculations. The software that implements the proposed methods is developed. The developed software is used in the analysis of metallographic images of materials for medical application.

Keywords: microstructure, image, processing automation, cluster, recognition

1 Introduction

Many devices for determining the medical parameters of a person represent information in graphical form. In particular, these are the results of studies using microscopes, the results of X-ray analysis, MRI, ultrasound, etc. In addition, medical instruments, implants, materials have a certain internal structure and surface structure, which must be controlled. Quality control in medicine plays an important role in ensuring the reliability of manufacturing prostheses, implants, medical instruments. Therefore, to obtain objective quantitative data in the processing of any information presented in graphical form, automation of image processing is required. Determining the quantitative characteristics of image elements is necessary for quick and correct interpretation of results. However, processing and analysis of medical, metallographic images is a very nontrivial task due to the complex shape and mutual arrangement of the elements, as well as the quality of the images provided by modern equipment.

Currently, effective methods of digital image processing have been developed [1, 2], including for the analysis of microstructures, surface coatings of products [3], etc. The complete process of image processing in analyzing the results of medical diagnostics and quality control of products is a complex multi-step time-consuming procedure and includes a number of basic steps: pre-processing and image restoration, segmentation, filtering, normalization of selected objects, recognition and, often, comparison with reference objects. Multistep is due to the fact that various processing tasks are closely related and are performed sequentially, therefore the quality of the solution of one of them affects the choice of the method of solving the other problems [4]. Automation of image processing is the main problem in obtaining quantitative information about the results of the study.

The aim of the study is to reduce the complexity of the process of automated recognition and obtaining a quantitative assessment of the parameters of graphic objects in medical images.

2 Problem statement

To implement automated image processing for medical purposes, it is necessary to solve the following tasks:

- the analysis of methods, models and information technologies for recognition of graphic objects in images, microstructures of implants and instruments;
- the development of a method for reducing the amount of information analyzed through the use of a specified dimensions grid, which is superimposed on the image;
- the development of an algorithm for recognizing objects in an image using software implementation for automating successive stages of recognizing objects in an image, including clustering methods, binarization methods, building concave contours, and applying these methods to processing medical images;
- the experimental obtaining of quantitative characteristics of recognized images and their elements on the basis of the developed methods, algorithms and software.

3 Literature review

Currently, the development of computing technology has significantly increased the resolution of microscopes, in addition, it has become possible to share a microscope with a computer, so it is possible to create software systems for analyzing and processing metallographic images. To obtain image characteristics by means of modern information technologies, a number of algorithms are used, which can be divided into three categories: algorithms based on comparison with templates; algorithms that use the methods of decision-making theory; algorithms using neural networks, etc. Each of these groups of algorithms has its own strengths and weaknesses, different areas of application. Methods using patterns are inflexible, but because of their simplicity and low resource intensity, they often are used in solving particular recognition problems. The method of neural networks, which uses the principles of artificial intelligence, is more versatile, but requires a significant amount of input data and computational resources necessary for the organization of the learning process; its disadvantage is poor predictability of recognition results. Algorithms based on methods of decision theory, make the learning process simple and provide an opportunity to achieve a high degree of predictability of recognition results.

Among the existing solutions in the field of analysis and processing of metallographic images are the following specialized software systems. Image Expert Pro 3 allows you to get a wide range of geometrical parameters of the elements of the microstructure. The resulting characteristics are available both for each type of microparticles separately, and in the form of their statistical sampling [5]. The Fuzzy Clustering and Data Analysis Toolbox is a Matlab-based software package that includes three main categories of functions: clustering algorithms (K-means, K-medoid, FCMclust, GKclust and GGclust), analysis functions (Dunn, Alternative Dunn, Xie and Beni's, Partition Index), as well as data visualization functions (Sammon's method) [6]. Altami Studio allows real-time image capture from various devices (microscope or camera), has the ability to pre-process bitmap images (rotate, crop, change brightness, gamma, contrast). Altami Studio provides the ability to generate analysis reports in a user-friendly format [7]. The authors have developed algorithmic software for a specialized software package that implements the use of clustering methods for processing metallographic images [8].

It should be noted that there is no universal mathematical apparatus that allows you to create a common formalized approach to the construction of systems for analyzing metallographic images. Therefore, a systematic approach is needed to solve practical problems of image processing, and the development of interactive tools to automate the process of obtaining quantitative information in metallographic studies. Existing software systems for solving problems of processing and analysis of metallographic images use various methods of cluster analysis. The principle of the K-means method is to build k clusters located at the greatest distances from each other. The number of required clusters is specified manually by the user [9]. The EM algorithm searches for clusters based on hidden characteristics that the algorithm determines based on known data about each object. The main disadvantage of the algorithm can be called probable quasi-optimal solutions for its work [10]. The C-means method is a modification of K-means, but taking into account fuzzy clustering, i.e. objects are not strictly belong to any cluster, but have a certain value of belonging to each cluster [11]. Kohonen neural networks are a two-layer neural network in which the neurons of the output layer are cluster elements corresponding to the number of clusters into which the given objects are divided. One of the drawbacks of the method is the need to pre-tune and train the neural network [12].

4 The method of "overlaying the grid" in image processing

To solve the set tasks, the authors analyzed the characteristics of metallographic images, identified the problems of their processing during process automation, associated with the need to analyze a large number of pixels in the presence of actual and process noise in the images. The authors propose a method and an algorithm for extracting image regions using cluster analysis [13] and grids, which will make it possible to abandon the processing of each image pixel.

Image analysis showed the redundancy of information on the image associated, for example, with the presence of the background on which the structure elements are located. The essence of the proposal is to reduce the amount of information analyzed, to use a grid of specified dimensions, which is superimposed on the image, and to analyze the image directly by grid pixels. The grid is built on the basis of the grid dimension entered by the user, which can be changed to ensure the best fit to the size of the elements selected in the image. Thus, the accuracy of determining the size of objects is regulated.

The next processing step is the selection of the desired points (pixels) in the image using a binarization algorithm. The binarization method implemented in this paper consists in comparing pixels in the RGB color model. Each pixel of the grid and the corresponding pixel on the image with a user-specified tolerance are compared. As a result of binarization, selected image pixels are obtained for further analysis.

Pseudocode of the proposed algorithm:

```
if (grid.R >= image.R – sens && grid.R <= image.R + sens)
if (grid.R >= image.R – sens && grid.R <= image.R + sens)
if (grid.R >= image.R – sens && grid.R <= image.R + sens)
return true;
```

where grid.R, grid.G, grid.B – a specified color in RGB format, image.R, image.G, image.B – RGB image color, sens – tolerance color parameters.

The algorithm was tested on a set of test color and grayscale images. To illustrate the operation, the result of the image binarization of the submicrocrystalline titanium BT 1-0 structure, obtained using transmission electron microscopy [14] and a color image obtained using an optical microscope is shown in Fig. 1. The search was performed according to the color of the object and the background color, with the use of the different grids sizes. An HSL color scheme has been added to analyze the image in grayscale. Each pixel was converted to HSL using standard formulas. Image analysis was performed within the grid.

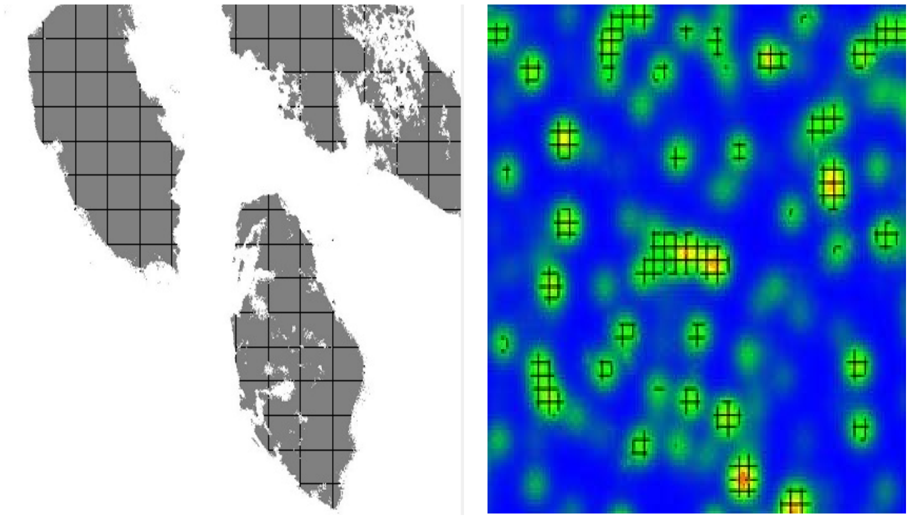
The results of the algorithm are obtained in the form of data arrays:

$$O^k = \{x_i, y_i, R_i, G_i, B_i, H_i, S_i, L_i, i = 1,2,3,...\}^k, k = 1,2,3,...,$$

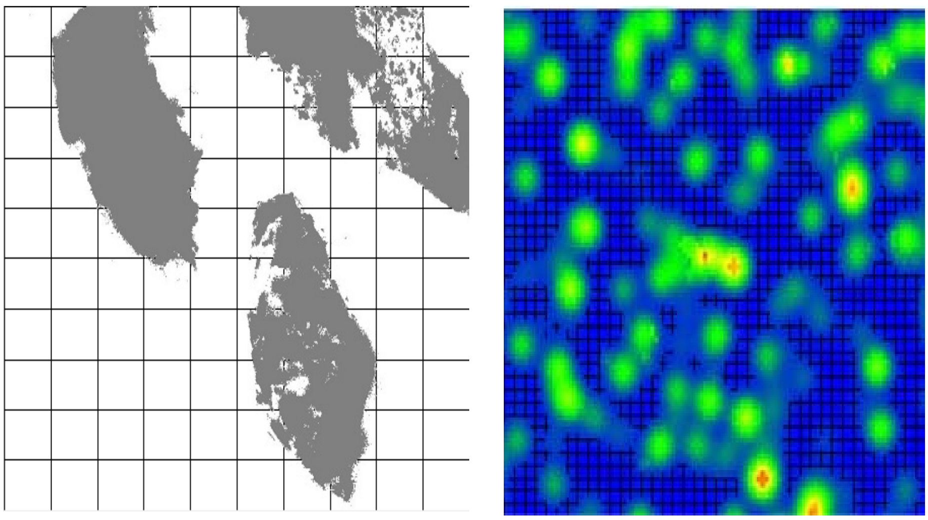
where k – the number of separate objects in the image; x_i, y_i – screen coordinates of the grid pixel; $R_i, G_i, B_i, H_i, S_i, L$ – RGB and HSL data for this pixel.

A simple contour can be constructed by connecting the “extreme” points of the grid superimposed on an object with the help of line segments. Analysis of the results of applying the algorithm on a set of test images showed that the amount of information processed is reduced by 4 – 25 times. This depends on the parameters of the elements in the image under study.

An important problem is the study of the effect of grid size on the quality of image processing and the exclusion of small elements of the original image. It is proposed to take the initial grid cell size no more than $d_x/5 \times d_y/5$, where d_x, d_y – the dimensions of the minimum picture element along the respective axes are. At this stage of work, the user for the whole image sets the grid size. The planned direction of the improvement of the algorithm is the splitting of the original image into separate zones.



a



b

Fig. 1. Selection of areas on the image using the “grid overlay” method with search by object color (a) and background color (b)

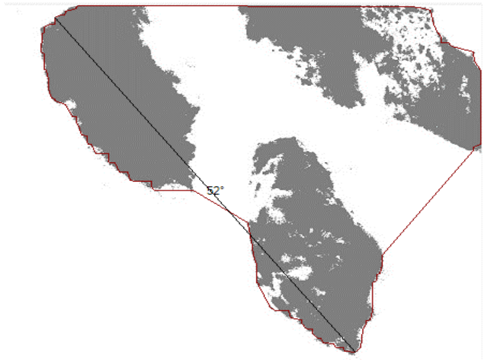
The resulting image points on the grid were used for the subsequent clustering algorithm. The algorithm was developed on the basis of the following requirements: the method should automatically determine the number of clusters based on the data provided; each analysis method must give the same result with the same input data. To find the clusters, you need to specify the value of the Euclidean distance $(dist)$ between the points inside the cluster. In this way, we can isolate the desired regions with different values of $dist$ to control the results of cluster allocation. The use of the grid in the considered algorithm allows us to significantly speed up the process of extracting the desired image elements.

The next task of analyzing the selected image elements associated with obtaining quantitative characteristics is to determine their orientation relative to the horizontal axis. To do this, build a segment, the extreme points of which will be the most distant grid points p_1, p_2 in the cluster, then the angle of the selected image areas is determined. To determine the length, the square of the Euclidean distance $d^2 = (x.p_2 - x.p_1)^2 + (y.p_2 - y.p_1)^2$ is used, which optimizes the execution time of the analysis. The result of applying the algorithm analyzing the test image is shown in Fig. 2.

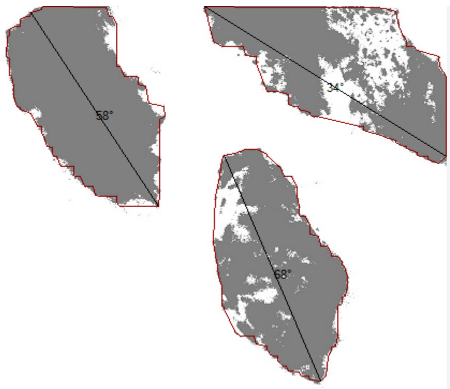
The next step to identify the elements in the image is to find the contour of each selected separate area. The problem of finding the convex contour of a set of points is well known and studied in detail. To solve it, there are many algorithms with different efficiency. The quick hull algorithm uses the idea of quick sort. The complexity of this algorithm in the average case is $O(N \cdot \log(N))$, where n is the number of starting points [15, 16]. Another well-known algorithm is Graham's three-step algorithm [17]. In the case of its implementation using quick sort, we obtain the total complexity of the algorithm $O(N \cdot \log(N))$. This also applies to the multi-step Jarvis algorithm [18]. The Divide-and-Conquer algorithm is a popular method for solving problems in computer science, it consists in dividing a problem into several parts, recursively looking for solutions in each of them and further integrating solutions to these problems into a common, complex solution. The obvious advantage of this approach is the possibility of paralleling the search for solutions to each individual task. A feature of the algorithm is that the algorithm for constructing contours itself can be performed using other, more optimal algorithms, with small volumes of initial points [19].

To solve the problem, an algorithm for constructing a concave contour was used. The idea of the algorithm is to first calculate the convex contour using any known algorithm, and then turn the convex contour into a concave [20]. The requirements for the algorithm are as follows: the algorithm must be stable; should be able to handle any set of points in size; the method should ensure the continuity of the contours, the absence of intersections between the contours and ensure that all points of the cluster fall into the selected contour. The speed of the algorithm of this part of the overall process is less important, given the basic tasks. In addition, from the user's point of view, the algorithm should be easy to use, allow control of the image processing process, the user is not required to enter a large number of additional parameters.

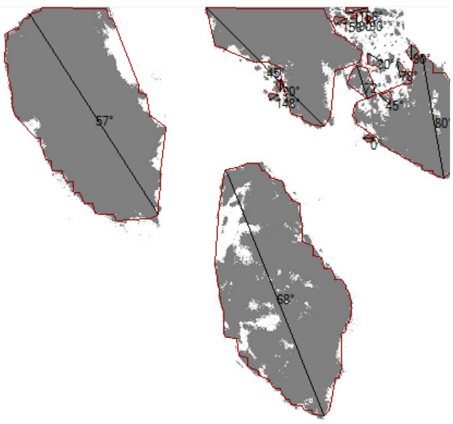
A concave contour of a set of points can be defined as a shape that minimizes the area of a figure bounded by a contour, but allows any angle between adjacent edges of the contour.



a



b



c

Fig. 2. Determination of clusters and the angle of inclination of the cluster to the horizontal axis at $dist = 70$ (a), $dist = 10$ (b), $dist = 2$ (c)

In most cases, there is no need to find the minimum area, as this can distort the figure. In our implementation of the concave contour method, we need a contour that defines and describes the general shape of a set of points. It is necessary that the contour is sufficiently smooth and show as much detail as possible.

The developed algorithm for calculating the concave contour first calculates the convex contour. The algorithm used in this case for calculating convex contours is the Divide and Conquer algorithm. Then the convex contour opens iteratively to create a concave contour. The first step of the concave contour construction algorithm is to add all the edges of the convex contour to the list, where they are sorted by length. Then the longest edge in the list is selected and removed.

The next step is to search through all the points that are most suitable for choosing the best candidate for adding a new part of the contour between the endpoints of the remote edge. The search criteria for points is that the largest angle from the new point to the end points of the old contour must be less than the cosine of the user-entered angle between the new edges (concavity coefficient K). This ensures that the point is not outside the contour. After each iteration, a check is made to see if the new edges of the contour intersect with the existing edges.

In the case of $K = 1$, we obtain an actually convex contour, the constant $K = 0$ gives a concave smooth contour, and the constant $K = -1$ gives a broken concave contour (Fig. 3). The ability to select the angle between new edges and a region-defined grid size to search for new points is an opportunity for the user to balance between a smoother or sharper contour.

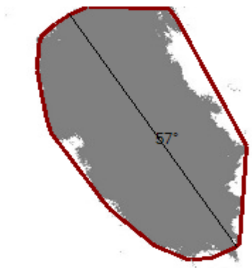
5 Conclusion

The analysis of the possibilities and shortcomings of existing medical image processing methods showed that in order to reduce the complexity of processing, it is necessary to automate the entire sequence of analysis steps with expanding the process control capabilities at each stage and using methods to reduce the amount of information processed.

An algorithm for identifying image element search areas and an algorithm for forming the boundaries of found clusters with concave sections with simultaneous application of the grid overlay method to reduce the amount of information processed was proposed and implemented.

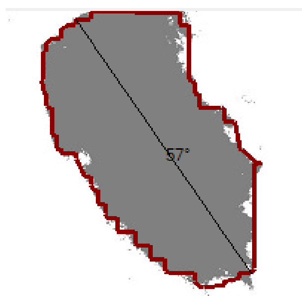
The object-oriented software that automates the stages of image processing is developed. The final result of the analysis is the areas highlighted in the image with concave contours and certain quantitative characteristics of their form.

The application of the grid overlay method made it possible to abandon the processing of each pixel of the image and reduce the processing time of the original data. The analysis of the results of the algorithm on the test image of the submicrocrystalline structure of titanium BT 1-0 showed that, depending on the problem being solved, it is possible to reduce the amount of information processed by 4-25 times.



a

b



c

Fig. 3. Examples of the construction of a convex contour (a), a smooth concave contour (b), a broken concave contour (c)

The prospects for further research are to test the proposed methods and algorithms on a wider set of applied problems, to study the dependence of the accuracy of working methods on the type of image.

References

1. Gonzalez, Rafael C., and Richard E. Woods. "Digital image processing." (2002).
2. Pratt, W. K. Introduction to digital image processing. CRC Press. (2013).
3. Panteleev, V. G., Egorova, O. V. and Klykova, E. I., (2005), *Computer microscopy*, [Komp'juternaja mikroskopija], Technosphere, Moscow.
4. Jaroslavskij, L. P., (1987), *Digital signal processing in optics and holography: introduction to digital optics*, [Cifrovaja obrabotka signalov v optike i golografii: vvedenie v cifrovuju optiku], Radio i svjaz', Moscow, 296 p.
5. NEXSYS ImageExpert Pro 3. http://www.nexsys.ru/nexsys_iepro3x.htm.
6. Balasko, Balazs, Janos Abonyi, and Balazs Feil. "Fuzzy clustering and data analysis toolbox." *Department of Process Engineering, University of Veszprem, Veszprem* (2005).
7. Altami Studio. http://alelso.ru/projects/altami-studio/.
8. Vasilyeva, L., Tarasov, A. and Getman, I. (2016), "Development of algorithmic support and model of the software complex of image processing", [Razrabotka algoritmicheskogo obespechenija i modeli programmnogo kompleksa obrabotki izobrazhenij], *Proceedings of the tenth international scientific-practical conference «Internet-Education-Science» (IES-2016)*, Vinnytsia, pp. 214–215, available at: http://ir.lib.vntu.edu.ua/handle/123456789/13371
9. Fasulo, D. (1999). *An analysis of recent work on clustering algorithms* (No. 01-03, p. 02). Technical report.
10. Oreshkov. V. (2012). EM-masshtabiruyemyy algoritm klasterizatsii. https://basegroup.ru/community/articles/em.
11. Leonenkov, A.V. (2003), *Fuzzy modeling in the MATLAB and fuzzyTECH environment* [Nechetkoe modelirovanie v srede MATLAB i fuzzyTECH], BHV-Peterburg, Peterburg, p. 385-388.
12. Friedman, J., Hastie, T., & Tibshirani, R. (2001). *The elements of statistical learning* (Vol. 1, No. 10). New York: Springer series in statistics.
13. Tarasov, O., Vasilyeva, L., & Efremov, M. (2017). Automation of processing of microstructures of metals based on contour and texture analysis of images. Naukovi Praci Donec'kogo Nacional'nogo Tehnicnogo Universitetu. Seria, Informatika, Kibernetika i Obcisluval'na Tehnika, 2(25), 109–117. doi:10.31474/1996-1588-2017-2-25-109-117
14. Tarasov A.F. et al (2015). *Osobennosti ispolzovaniya protsessa reversivnogo sdviga dlya polucheniya submiktrokristallicheskikh obyemnykh zagotovok* [Features of using the reverse shift process for obtaining submicrocrystalline bulk billets]
15. Barber, C. B., Dobkin, D. P., Dobkin, D. P., & Huhdanpaa, H. (1996). The quickhull algorithm for convex hulls. *ACM Transactions on Mathematical Software (TOMS)*, *22*(4), 469-483.
16. Wenger, R. (1997). Randomized quickhull. *Algorithmica*, *17*(3), 322-329. https://doi.org/10.1007/BF02523195
17. The Graham Scan, Convex Hull Algorithm. http://www.algomation.com/algorithm/graham-scan-convex-hull.
18. Jarvis, R. A. (1973). On the identification of the convex hull of a finite set of points in the plane. *Information processing letters*, *2*(1), 18-21. https://doi.org/10.1016/0020-0190(73)90020-3
19. Bentley, J. L. (1980). Multidimensional divide-and-conquer. *Communications of the ACM*, *23*(4), 214-229.
20. Lee, D.T. & Schachter, B.J. International Journal of Computer and Information Sciences (1980) 9: 219. https://doi.org/10.1007/BF00977785