A Technique for Extraction of Diagnostic Data from Cytological Specimens*

Igor Gurevich, Andrey Khilkov, and Dmitry Murashov

Scientific Council "Cybernetics" of the Russian Academy of Sciences. 40, Vavilov street, Moscow, GSP-1, 119991, Russian Federation
igourevi@ccas.ru,dmmur@rbcmail.ru

Abstract. In this paper, a possibility of developing a new criterion for diagnostics of hematopoietic tumors, such as chronic B-cell lymphatic leukemia, transformation of chronic B-cell lymphatic leukemia into lymphosarcoma, and primary B-cell lymphosarcoma, from images of cell nuclei of lymphatic nodes is considered. A method for image analysis of lymphatic node specimens is developed on the basis of the scale space approach. A diagnostically important criterion is defined as a total amount of points of spatial intensity extrema in the families of blurred images generated by the given image of a cell nucleus. The procedure for calculating criterion values is presented.

1 Introduction

A large quantity of research in image processing and analysis are directed at the development of medical diagnostics. Recently appeared a new perspective trend concerned with the development of diagnostic techniques for automated analysis of morphology of blood cells and hematopoietic organs using analysis of microscopic images. In this paper, a relatively small sample of images is used for obtaining the criterion for diagnostics of hematopoietic tumors, such as chronic B-cell lymphatic leukemia, its transformation to lymphosarcoma, and primary B-cell lymphosarcoma (according to the classification of A. Vorob'ev and M. Brilliant [5]).

Experts in hematology have found out, that specimen cell nuclei of a tissue of lymphatic nodes taken from patients with the malignant tumor diagnose are larger than those taken from patients with the non-malignant tumor diagnose. Thus, an obvious diagnostic criterion is the area of cell nucleus. But this criterion is unsuitable for more accurate diagnostics: it is impossible to distinguish such diseases as transformation of chronic lymphoid leukemia and lymphosarcoma.

The procedure of searching for a diagnostic criterion includes the following steps: the experts indicate the diagnostically important cell nuclei in the images of lymphatic node specimens of three groups of patients having the diagnosed diseases. These images are considered as an input information. Next, the developed method of specimen

* This work was partly supported by the Russian Foundation for Basic Research (grants NN 01-07-90016, 01-07-90017, 03-07-90406).

A. Sanfeliu and J. Ruiz-Shulcloper (Eds.): CIARP 2003, LNCS 2905, pp. 274–281, 2003.
© Springer-Verlag Berlin Heidelberg 2003

image analysis is used for calculating qualitative characteristics (features) from the indicated nuclei. The obtained values are analyzed, and this gives an opportunity to formulate a criterion for making diagnostic decisions. The proposed method for specimen image analysis is based on the known scale space approach [1,3,4].

2 Properties of Cell Images and Requirements to the Method

The image of a lymphatic node specimen is a color image taken by a camera and enlarged by a microscope (24 bits per pixel). The size of the image is 1536x1024 pixels covering a site of 60—100 microns in diameter. The resolution is 0,06 microns per pixel. The analyzed objects are the fragments of the gray-scale specimen images containing cell nuclei. These images are characterized by inhomogeneous coloring and by the presence of dark spots and bright areas representing their internal structure.

For a diagnostics, experts pay a special attention to the cells of two classes: mature cells, with the mature structure of chromatin, (see Fig. 1) and sarcoma cells, with the immature structure of chromatin, (see Fig. 2) [5-7]. In the first case (chronic lymphatic leukemia), with few exceptions, the image contains only mature cells. In the case of sarcoma transformation of the chronic lymphatic leukemia, the specimen contains both mature and immature (sarcoma) cells. In the case of primary lymphatic sarcoma, the sarcoma cells prevail in the image.

It is necessary to take into account such specific properties of cell images as low dye quality, instability of specimen characteristics, non-uniformity of specimen light exposure during microscoping, presence of damaged and unsuitable for analysis cells. Mature chromatin is homogeneous with light furrows. Immature chromatin can have a filamentous structure of different patterns, a fibrous, or a granular structure [7]. The analysis of cell nucleus images should yield quantitative characteristics that capture the structure and pattern of chromatin.

In view of the specific properties of cell images listed above, the following requirements to the method are formulated: (a) suitability for selection of features for classification of cell images; (b) resistance to noise caused by image acquisition process and specimen quality; (c) resistance to errors and noise of image processing algorithms; (d) correspondence of classification results to expert estimations.

The quantitative analysis cytological and tissue specimen images is based on the evaluation of shape, intensity, and textural features. In practice, the great attention is paid to automated analysis of chromatin arrangement in the cell nuclei. It has been proven in many studies that chromatin distribution corresponds to the state of malignancy. Two basic approaches to analysis of a chromatin constitution are known [9]. Within the first, structural, the chromatin distribution is considered as a local arrangement of rather small objects of varying intensity. The intensity features of dark and bright particles are evaluated. This approach is substantially heuristic. The second approach, textural, is based on the statistical characteristics of chromatin arrangement and related to analysis of the regularities of chromatin structure. Applied in practice methods for textural analysis use grey level dependency matrices [8], co-occurrence, run-length features, rice-field operators, and watersheds (topological methods) [10],

Fig. 1. Grayscale images of mature cell tumor

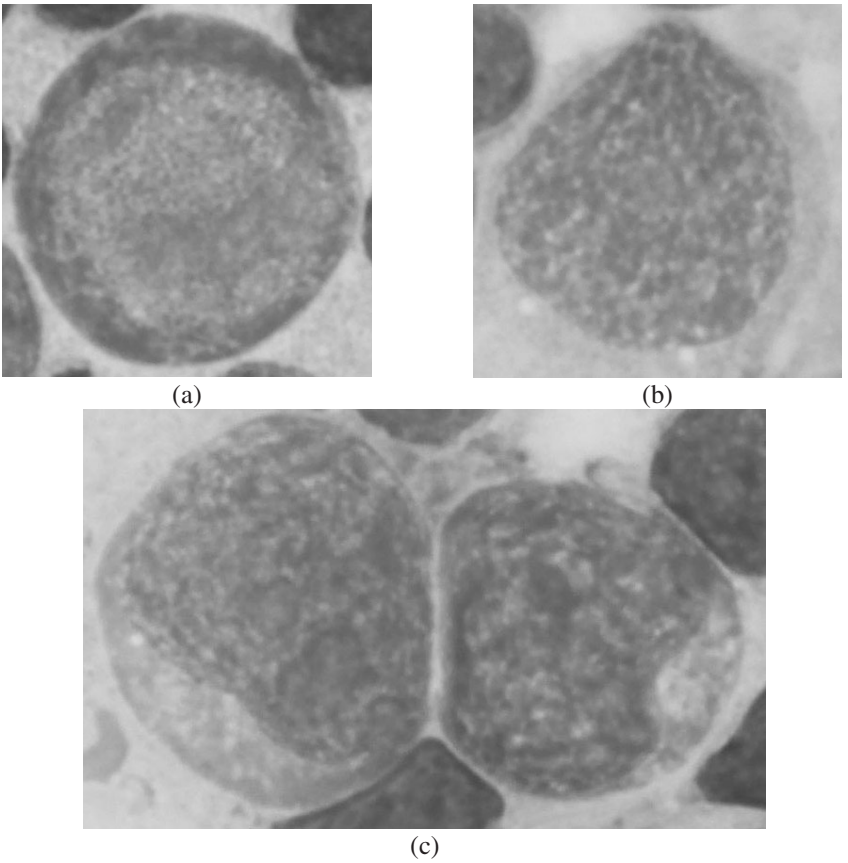


Fig. 2. Grayscale images of sarcoma cells of lymphatic nodes: filamentous structure of chromatin (a); granular (b); fibrous (c)

heterogeneity, clumpiness, margination, and radius of particles [12] (the Mayall/Young features), invariant features (polinomial invariants).

The main disadvantage of known textural methods [9] is their sensitivity parameters and conditions of image acquisition, to properties of researched preparations, and also to precision of microscope focusing.

Below, we consider a method for analysis of cell nucleus images that was used for searching for a diagnostic criterion. The proposed method combines features of both approaches: on the one hand the intensity features of the chromatin particles are analyzed, and on the other hand, the diagnostic criterion is formulated in terms of the simple quantitative characteristics, describing the chromatin structure of cell nuclei – the amount of intensity extrema in the families of blurred cell nuclei images. This feature is related to the amount of chromatin particles and characterizes the state of malignancy.

3 Method for Analysis of Nuclei Images

Among contemporary approaches to image analysis the approach of Gaussian scale space entirely meets the listed above requirements [1,3,4]. The scale space technique provides properties of invariance with respect to shift, rotation, scaling, and linear transformations of intensity. It decreases the sensitivity of the analysis to microscope focusing. The concept of the scale space gives the natural way to represent an input image $L(\mathrm{x})$ ($L(\mathrm{x})$ is the intensity function of spatial coordinates) at finite resolution by convolving it with a Gaussian kernel $G(\mathrm{x},t)$ of various widths, thus obtaining a smoothed image at a scale determined by the width $\sigma = \sqrt{2t}$ (t – is a scale parameter). $L(\mathrm{x},t)$ satisfies the heat equation. The heat equation generates a family of blurred images [3]. As t increasing the blur effect grows, and fine details of the image are lost. The properties of constructed scale space reflect properties of the initial image. Scale space properties are explored by using localization of its critical points.

The proposed method for cytological specimen analysis consists in construction of a family of blurred images (scale space) for various t and selection of diagnostic criterion using localization of scale space critical points. Critical points reflect the internal structure of the objects in the image, and exploration of the entire family of derived images allows one to analyze both fine details and large structural elements.

3.1 Main Objectives of the Proposed Method

Taking into account the concept of scale space approach, the following problems should be solved during cell image analysis and selection of diagnostic criterion: (a) construction of the one-parameter family of derived images $L(\mathrm{x};t)$ (a scale space) from initial image $L(\mathrm{x})$ for different diseases; (b) extraction of critical points in scale space images; (c) analysis of spatial critical points distribution for different groups of patients for diagnostic criterion selection; (d) calculation of diagnostic criterion values.

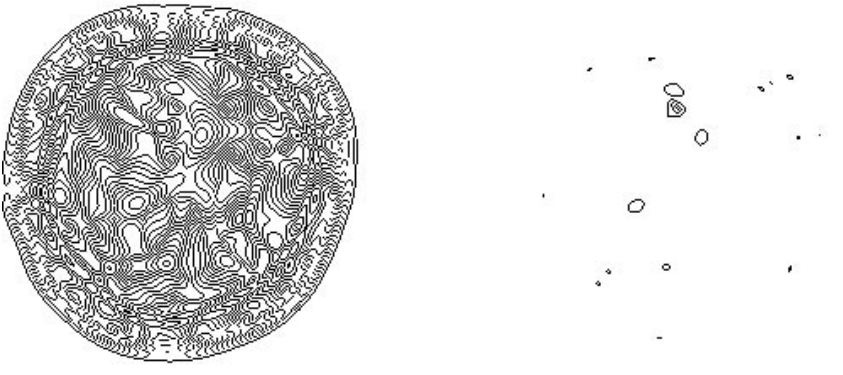
3.2 Construction of a Family of Blurred Images

For construction of a family of blurred images, it is necessary to determine the range and the step of scale parameter t. The computational experiments have shown that the analysis of critical points in image family is expedient for the values of scale parame-

ter in the range of $8 \leq t \leq 60$. The step value of a scale parameter should be taken in the range of $0.005 \leq \Delta t \leq 0.032$. As a result, the families of blurred images corresponding to the scale spaces of malignant and non-malignant cell nuclei images were constructed.

3.3 Localization of Spatial Critical Points

According to the scale space approach, the derived families of images were explored for detection of critical points.

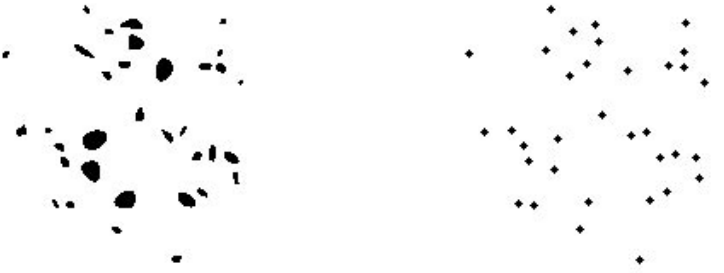


(a) (b)

Fig. 3. Spatial gradients (L_{x1} (x;t) OR L_{x2} (x;t)) at t = 32 (a) and extracted closed curves around extrema in a single scale-space image (b) (negative images)

For localization of critical points within proposed method, the topological properties of iso-intensity manifolds in the neighborhoods of critical points [1] are used. The algorithm for selection of closed loops (curves of a nonzero gradient values, bounding iso-intensity curves around points of extremum) is applied. A special procedure that includes the standard image processing operations was developed. It consists in the following steps:

1. The following operations are carried out for each image in the family: (a) logical summation of images $L_{x_1}(x;t)$ and $L_{x_2}(x;t)$; (b) thresholding of the resulting image; (c) removing of the "rubbish"; (d) overlaying of a nucleus mask to restrict the area of interest and remove the residual noise at the peripheries of a nucleus.
2. All scale space images processed at Step 1 are overlaid (logical OR).
3. The morphological operations are applied in order to fill regions bounded by closed curves of nonzero gradient and to remove residual rubbish.
4. The coordinates of the geometrical centers of the filled regions (the neighborhoods of extrema) are found.
5. The total amount of the centers of the filled regions is calculated.



(a) (b)

Fig. 4. Neighborhoods of scale space spatial extrema (a), centers of extrema neighborhoods (b)

In Figs. 3, 4, the steps of extrema localization are illustrated. In Fig. 3 (a), the logical "OR" of spatial gradient images at t=32 is presented. In Fig. 3 (b), one can see the extracted closed curves around extrema at t=32. In Fig. 4 (a), the neighborhoods of extrema for the whole family of scale space images are filled with black and, in Fig. 4 (b), the centers of colored areas are presented.

3.4 Selection of Diagnostic Criterion

The developed procedure was applied to the analysis of scale spaces, generated by images of lymphatic node specimens for diagnoses of malignant (primary B-cell lymphosarcoma and transformation of chronic B-cell lymphocytic leukemia) and non-malignant (chronic B-cell lymphocytic leukemia) tumours. 86 images of cell nuclei from 25 patients were analyzed. Families of blurred images for scale parameter in the range of $12.5 \le t \le 50$ were generated and explored. In Table 1 the statistical characteristics of amount of spatial extrema for various diagnoses are given.

Using the results of experiments, the chart displaying the characteristics of cell nuclei images, such as total amount of spatial extrema n and the area of a nucleus s, was created (see Fig. 4). The chart area (see Fig. 4) includes three significant parts: (I) the area, located to the left of value s = 137 and below n = 60; (II) the area, where $137 < s < 200$ and $n > 60$; (III) the area to the right of value s = 200.

The first area mainly contains the points corresponding to the diagnose of chronic lymphocytic leukemia (CLL). In the second area, the transformation of chronic lymphocytic leukemia (TRCLL) is dominating. The third area contains transformation of chronic lymphocytic leukemia as well as lymphosarcoma (LS). For classification of cell nuclei located in area (III), it is possible to construct a separating functions. The spread of points in Fig. 4 in the region (III) is caused by the different types of structure and pattern of chromatin of the malignant cell nuclei (see Fig. 2). Therefore, the more accurate classification requires analysis of critical points for different types of chromatin structure.

Table 1. Range of spatial extrema total amount for specimen images corresponding to various diagnoses

Diagnose	Min	Max
Chronic lymphocytic leukemia	23	60
Transformation of chronic lymphocytic leukemia	44	216
Lymphosarcoma	59	175
Transformation of chronic lymphocytic leukemia and lymphosarcoma	44	216

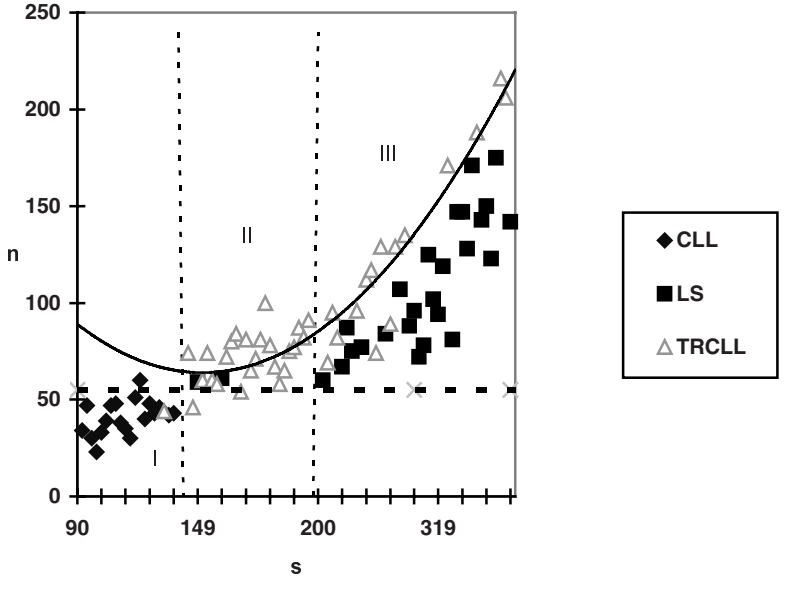


Fig. 5. Distribution of cell nuclei in coordinates "nucleus area","amount of extrema" (s,n)"

The results presented in Fig. 4 and in Table 1 allow us to conclude that the total amount of spatial extrema in cell nuclei images may be used as a diagnostic criterion. A special technique for calculation of the diagnostic criterion value is developed and implemented in the "Black Square" software system [2].

4 Conclusions and Directions of Further Research

We considered a possibility of developing a new criterion for diagnostics of hematopoietic tumors from the images of cell nuclei of lymphatic nodes. The results are as follows.

1. The method for analysis of the images of lymphatic node specimens is developed.
2. A diagnostically important criterion is obtained; it is defined as a total amount of of spatial extrema in scale space generated by the image of a cell nucleus.

3. The technique for calculating the diagnostic criterion value is developed and integrated in the "Black Square" [2] system library.

The further research will be aimed at (a) increasing the precision of critical points localization, (b) selecting of diagnostic criteria based on the analysis of all types of critical points, and their evolution at the increasing scale parameter; (c) augmenting the sample of cell images. At the final stage of research, the decision rules for making diagnostic decisions will be formulated.

Acknowledgements. We are grateful to Prof. I.A. Vorob'ev, Dr. D.V. Kharazishvili, and their colleagues in the Hematological Center of the Russian Academy of Medical Sciences for providing a material for this work and for their professional advices.

References

1. Florack, L.M.J., Kuijper, A.: The topological structure of scale-space images. Journal of Mathematical Imaging and Vision, Vol. 12, N1 (2000) 65–80
2. Gurevich, I.B., Khilkov, A.V., Murashov, D.M., Smetanin, Yu.G., Zhuravlev, Yu.I.: Black Square Version 1.0: Programm Development System for Automation of Scientific Research and Education. Pattern Recognition and Image Analysis: Advances in Mathematical Theory and Applications, Vol. 9, N4 (1999) 609–634
3. Koenderink, J.J.: The structure of images. Biol. Cybern., Vol.50. (1984) 363–370
4. Lindeberg, T.: Scale-space Theory in Computer Vision. The Kluwer International Series in Engineering and Computer Science. Kluwer Academic Publishers (1994)
5. Vorob'ev A.I., ed. Atlas "Tumors of lymphatic system", Hematological Scientific Center of the Russian Academy of Medical Sciences (2001)
6. Vorob'ev A.I., Vorob'ev I.A., Kharazishvili D.V., et.al. Recognition of lymphatic tumours of a mediastinum. Problems of a tuberculosis, 2 (2001) 21–25
7. Kharazishvili D.V., "Morphometrical analysis of intracellular neoplasms of lymphatic system", Hematological Scientific Center of the Russian Academy of Medical Sciences, Ph.D. thesis, (2001)
8. Weyn B., Van de Wouwer G., Koprowski M., Van Daele A., Dhaene K., Scheunders P., Jacob W., and Van marck E. Value of morphometry, texture analysis, densitometry and histometry in the differential diagnosis and prognosis of malignant mesothelioma, Journal of Pathology, 189 (1999) 581–589
9. K. Rodenacker, and E. Bengtsen. A feature set for cytometry on digitized microscopic images. Anal Cell Pathol, 25(1) (2003)1–36
10. K. Rodenacker. Applications of topology for evaluating pictorial structures. In Reinhard Klette and Walter G. Kropatsch, editors, Theoretical Foundations of Computer Vision, Akademie-Verlag, Berlin, (1993) 35–46
11. Rodenacker. Quantitative microscope image analysis for improved diagnosis and prognosis of tumours in pathology. In Creaso Info Medical Imaging, Vol. 22 (1995)
12. I.T. Young, P. Verbeek and B.H. Mayall, Characterization of chromatin distribution in cell nuclei, Cytometry 7(5) (1986) 467–474