CELL NUCLEI SEGMENTATION USING FUZZY LOGIC ENGINE

Grigory Begelman[1], *Eran Gur*[3], *Ehud Rivlin*[1], *Michael Rudzsky*[1], *and Zeev Zalevsky*[2]

[1]Computer Science Dept., Technion- Israel Institute of Technology 3200 Haifa, Israel
[2]School of Engineering, Bar-Ilan University, Ramat-Gan 52900, Israel
[3]Faculty of Eng., Tel-Aviv University, 69978 Tel-Aviv, Israel

ABSTRACT

The task of segmenting cell nuclei in microscope images is a classical image analysis problem. The accurate nuclei segmentation may contribute to development of successful system which automate the analysis of microscope images for pathology detection. In this article we describe a method for semi-supervised training of fuzzy logic engine. The fuzzy logic engine is applied to connect a set of parameters proven to be important for nucleus segmentation. In addition each parameter for itself is detected using a set of fuzzy logic rules. We present results of nuclei segmentation using fuzzy logic set of rules.

1. INTRODUCTION

The fuzzy logic (FL) control was first introduced by Zadeh [1] in 1965. The difference between fuzzy logic control and classical control is that in a classical case data either belongs to a set or not. In FL the degree of belonging to a fuzzy set is called membership grade (MG) that depends on the structure of the fuzzy set described by a membership function (MF).

Fuzzy logic-based rule induction can handle noise and uncertainty in data values well. Selecting and extracting useful attributes of target objects in noisy data is not an obvious task. In these case, domain knowledge and user analysis can be useful. Fuzzy logic based models utilize the domain knowledge in coming up with rules of data selection and extraction. In addition, it possible to implement the fuzzy logic rules as an optical system thus obtaining real-time segmentation [2].

Automatic nucleus segmentation is a difficult task: nuclei vary in color and shape. Moreover, they are surrounded by cytoplasm having similar color attributes. In order to distinguish the nuclei from the background we have used the following features: color, circularity, and object dimensions. A statistical model was derived for the distribution of these features and the fuzzy logic engine was trained according to the distributions of the features. Then the features were combined into fuzzy rules. We demonstrate the nuclei segmentation results based on the fuzzy logic engine that employs these rules.

The paper is organized as follows. Section 2 presents overview of previous work. Section 3 is a short introduction into fuzzy logic. Training process of the fuzzy logic engine in described in section 4. Experimental results are presented in section 5. Section 6 concludes the paper.

2. PREVIOUS WORK

Cell nuclei segmentation is an important stage in microscope image analysis. Unsupervised nucleus segmentation based on dual active contour algorithm was used in [3]. In [4] morphological watersheds was used for nuclei segmentation. The oversegmentation problem peculiar to this method was prevented by using an efficient homotopy image modification module. In [5] multistage segmentation algorithm was proposed. In the first stage Compact Hough Transform is used to determine possible locations of nuclei, then nuclei boundary was calculated by optimization of boundary likelihood function. Energy and entropy features extracted from multiwavelet coefficients of an image and co-occurrence matrix were used for grading of pathological images of prostate in [6]. General technique for the analysis of multimodal data based on the mean shift nonparametric analysis [7] was applied in [8] for cell image segmentation. Fuzzy k-nearest neighbor algorithm was used in [9] and[10] to assess the accuracy of oncological prognosis and to determine the significance of prognostic markers.

3. FUZZY SYSTEMS

The theory of fuzzy logic was developed by Zadeh in the late 1960s [1]. This was the rediscovering of the multivalued logic created by Lukasiewicz. The main idea was that most of the phenomena of the real world could not be described by two values, as is assumed in the classical set approach. A fuzzy set is a set without a crisp boundary. That is, the translation from "belonging to a set" to "not belonging to a set" is gradual, and this smooth transition is characterized by membership functions that give fuzzy sets

0-7803-8554-3/04/$20.00 ©2004 IEEE.

flexibility in modeling commonly used linguistic expression such as "the object is dark" or "the object is round".

The fuzzy sets and membership functions are defined in the following manner: if X is a collection of objects, then a fuzzy set A in X is defined as a set of ordered pairs:

$$A = \{(x, \mu_A(x))|x \in X\}.$$

In the above equation, A is a fuzzy set and $\mu_A(x)$ is a membership function (MF in short) of x in A. The MF maps each element of X to a continuous membership value (or membership grade) between 0 and 1. There are many types of membership functions (trapezoid, Gaussian, generalized bell, sigmoidal, etc.). For our application sigmoidal and Gaussian MF are chosen.

A fuzzy rule base contains fuzzy rules R_i:

$$R_i : IF(x_1 isA_{i1}) \wedge (x_2 isA_{i2}) \wedge \cdots \wedge (x_n isA_{in}) \\ THEN(y isB_i),$$

where A_{ij} and B_i are fuzzy sets, x_j and y are fuzzy inputs and output correspondingly. The structure of a rule is the following:

$$IF\ \ Premise\ \ THEN\ \ Conclusion,$$

where the premise consists of antecedents linked by fuzzy operator $\wedge$ (AND). There are many alternative possibilities to define the fuzzy operator. The most frequently used one is the minimum operation. In the MAX-MIN inference method, the activation degree of the premise is:

$$w_i = min\{w_{i1}, w_{i2}, \ldots, w_{in}\},$$

where w_{ij} is the intersection between input x_j and fuzzy set X_{ij} in the i^{th} rule. In a fuzzy inference system the output is calculated in the following way: the activation degree w_i is computed for each rule in the rule base. Then the i^{th} output fuzzy set is cut-off by w_i in each rule and the union of these cut-off fuzzy sets is composed, where the union means generally the maximum operation. Then a crisp value from the resultant output fuzzy set is calculated. This process is called defuzzification. There are many ways of defuzzification. The Center of Gravity (COG) method is used here because it is general and easy to compute. This method calculates the crisp output by the sums of the center of gravities of the conclusions. Thus, a fuzzy inference system can compute output y of an input vector x.

4. TRAINING THE FUZZY LOGIC ENGINE

Fuzzy logic engine training is a first important step in image segmentation. There are several reasons making the segmentation difficult. The first reason is a nonuniform illumination of the slide as a result of optical abberations (the intensity of central part of the captured images is greater than the intensity of the periphery area). The second reason is color variation between slides as a result of color fading of the stained tissues. Another reason is different illumination conditions of the images as they were captured in different time. In addition, the captured images contain various artifacts (e.g. dust).

In order to make more robust segmentation we take into account both color features and shape information of the nuclei.

4.1. Color features

Different illumination and color variation between slides make it impossible to perform segmentation of microscope objects based on precalculated color features, so the calculation of the color parameters is performed for each image individually.

Generally, there are three types of distinguishable objects in the images: muscular tissue (stroma), cells' nuclei, and prostate glands. We model the color distribution of the images as a mixture of Gaussian distributions, one Gaussian component for each type of object:

$$f(x|\xi) = \sum_{i=1}^{3} \alpha_i f_i(x|\xi_i), \qquad (1)$$

where x is a feature vector consisting of the color components in RGB color space, the α_i represents the mixing weights for each component of the mixture ($\sum_{i=1}^{3} \alpha_i = 1$), ξ represents the collection of parameters $(\alpha_i, \xi_i)_{i=1,2,3}$, and f_i is a multivariate Gaussian density parameterized by ξ_i. We use the Expectation-Maximization algorithm to [11] to

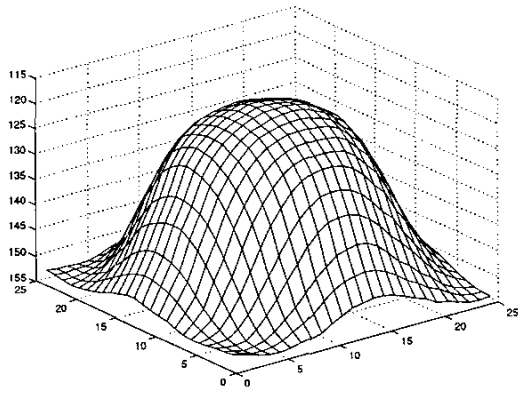

Fig. 1. Mean nucleon gray scale level.

determine the maximum likelihood parameters estimates of f in the feature space. Once the model's parameters are estimated, the components of Gaussian mixture are identified using the prior information about color relations of

the enumerated object types: glands have the highest intensity while the nuclei have stronger blue component than the stroma area. Finally we initialize the Gaussian member function with the estimated parameters.

4.2. Shape features

Color-based segmentation is not sufficient for nuclei detection as the microscopic images contain many non-nuclei objects having nuclei color. To achieve better segmentation we combine the color and appearance information of the nuclei.

To this end the nucleon template is calculated from a number of manually selected nuclei (see figure 1 for the example of mean nucleon). Then the nucleon template is matched against the image using normalized cross correlation (the matching is performed on gray-scale images). Thus high cross correlation value at particular image location is an additional evidence of nucleon presence.

5. EXPERIMENTAL RESULTS

Our database consists of microscopic color images of prostate tissue samples stained by hematoxylin and eosin. Nuclei

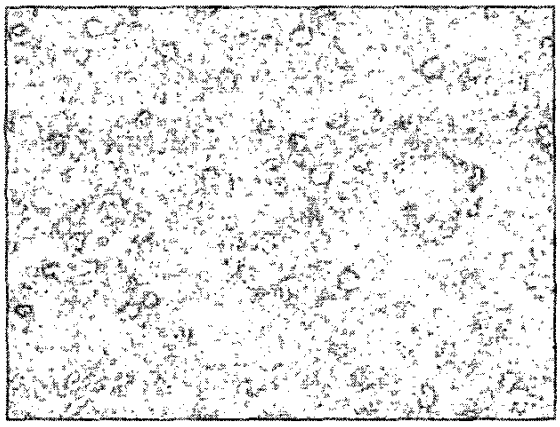

Fig. 2. Input sample.

of cells are stained by various blue hues (basophilic stain). They are surrounded by a pinkish background(eosinophilic stain).

In order to estimate color features, CIE Lab color space was chosen because it is a perceptually uniform color space, and the Euclidean distance between two colors in this space is perceptually uniform measure for color difference.

The mean nuclei was calculated by averaging 400 manually selected nuclei images (the dimensions of the picked nuclei images is 35×35 pixels). The area circularity was estimated by calculating normalized cross correlation of the gray scale image with the mean nuclei.

The set of fuzzy rules consisted of four terms:

1. If the shape is not circular, then this is a background area,
2. If the color is closer to the nucleon color than to the background color, then this is a nucleon area,
3. If the color is closer to the background color than to the nucleon color, then this is a background area,
4. If the shape is circular and the color of strongly blurred image is closer to the nucleon color than this is a nucleon area.

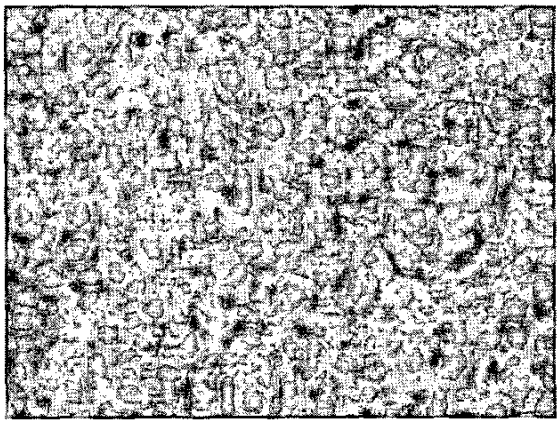

Fig. 3. Circularity.

A typical microscope histology image was chosen to demonstrate segmentation performance (figure 2). The calculated circularity feature is presented on the figure 3. Each pixel of the image was classified by the fuzzy logic engine according to the enumerated rules.

In order to estimate the segmentation performance we compare the final segmentation result (see figure 4) with segmentation provided by non-fuzzy color-based method (see figure 5).

The latter result was obtained by segmentation based on the color distribution estimation. The multivariate Gaussian mixture model was assumed for the color distribution (see equation 1). Then Expectation-Maximization algorithm was used for estimation of the mixture parameters. The mixture components are identified using prior color relations. Then we assign each pixel to one of three clusters according to the MAP estimate of point's membership based on its color features:

$$cluster\#(x) = \arg\max_i \alpha_i f_i(x|\xi_i)$$

The stroma and glands area are denoted by black color, while the nuclei area is denoted by white.

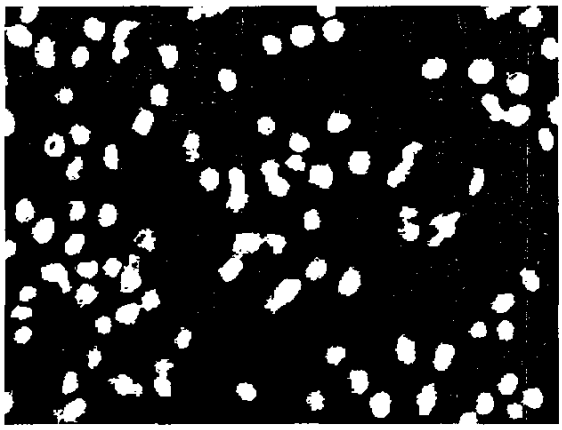

Fig. 4. Fuzzy segmentation result.

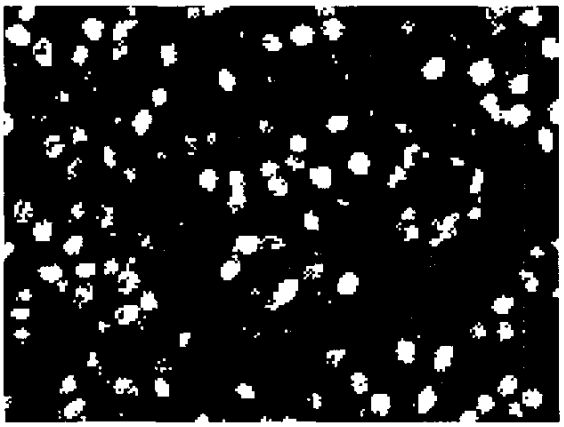

Fig. 5. Non-fuzzy color-based clustering.

6. CONCLUSIONS

In this paper we have demonstrated the usage of fuzzy classification engine in nuclei cell segmentation. The fuzzy rules were based on shape and color features. The classification engine was set up with statistically estimated distribution parameters of image features and verified on a large microscope image data set. The fuzzy method exhibited better segmentation results then segmentation based on crisp rules.

7. ACKNOWLEDGEMENTS

The authors would like to thank Edmond Sabo from Department of Pathology, Carmel Medical Center, Haifa, Israel for assistance in creating microscope image database.

8. REFERENCES

[1] L.A. Zadeh, "Fuzzy Sets," Inf. Conrol, 8(1965), 338-353.

[2] E. Gur, D. Mendlovic and Z. Zalevsky, "Optical implementation of fuzzy-logic controllers. Part I," Appl. Opt., 37, 1998, 6937–6945.

[3] P. Bamford, B. Lovell, "Unsupervised cell nucleus segmentation with active contours," Signal Processing, 71(1988), 203-213.

[4] J. A. Ferreira Costa, N. D. Mascarenhas, and M. L. de Andrade Netto, "Cell nuclei segmentation in noisy images using morphological watersheds," Proc. SPIE Vol. 3164, p. 314-324, 1997. Applications of Digital Image Processing XX.

[5] T. Mouroutis, S. J. Roberts, A. A. Bharath, "Robust Cell Nuclei Segmentation Using Statistical Modelling," Bioimaging, Vol 6(2), pp 79-91, 1998.

[6] Kourosh Jafari-Khouzani and Hamid Soltanian-Zadeh, "Multiwavelet grading of pathological images of prostate," *IEEE Transactions on Biomedical Engineering*, vol. 50, no. 6, pp. 697–704, June 2003.

[7] Y. Cheng, "Mean shift, mode seeking, and clustering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, no. 8, pp. 790–799, August 1995.

[8] D. Comaniciu and P. Meer, "Cell image segmentation for diagnostic pathology in advanced algorithmic approaches to medical image segmentation: State-of-the-art applications in cardiology, neurology, mammography and pathology," 2001.

[9] Dobrila Petrovic Raouf N. Gorgui-Naguib H. Seker, Michael O. Odetayo, "A fuzzy logic based-method for prognostic decision making in breast and prostate cancers," *IEEE Transactions on Information Technology in Biomedicine*, vol. 7, no. 2, pp. 114–122, 2003.

[10] H. Seker, M.O. Odetayo, D. Petrovic, R.N.G. Naguib, and F. Hamdy, "An intelligent hybrid neuro-fuzzy rule-based system for prognostic decision making in prostate cancer patients," in *Proc. of the 4th Annual IEEE Engineering in Medicine and Biology Society Special Topic Conference on Information Technology Applications in Biomedicine*, S. Becker, S. Thrun, and K. Obermayer, Eds., 2003.

[11] A. Dempster, N. Laird, D. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," J. Royal Statistical Soc., 39(1), 1977, 1–38.