MVA2007 IAPR Conference on Machine Vision Applications, May 16-18, 2007, Tokyo, JAPAN

Automatic Cell Image Segmentation using a Shape-Classification Model

Shishir Shah
University of Houston
Dept. of Computer Science
Houston, TX 77204, U.S.A.
shah@cs.uh.edu

Abstract

This paper presents a segmentation method for detecting cells in immunohistochemically stained cytological images. A two-phase approach to segmentation is used where an unsupervised clustering approach coupled with cluster merging based on a fitness function is used as the first phase to obtain a first approximation of the cell locations. A joint segmentation-classification approach incorporating ellipse as a shape model is used as the second phase to detect the final cell contour. Results of segmentation are presented and compared to ground truth measurements.

1. Introduction

Pathologists often make diagnostic decisions by observing cells and their spatial distributions in the biological specimen, in particular, the geometric parameters of the cell such as the area, radius, and the circumference [1]. Thus, in an automated system it is very useful to accurately measure the geometric parameters of the biological specimen. As a precursor to accurate measurement, segmentation of cells is required. Automatic cell segmentation is one of the most interesting segmentation problems due both to the complex nature of the cell tissues and to problems inherent to microscopy. Cytological images for immunohistochemical analysis, share the following characteristics:

- Poor contrast, i.e., cell gray levels may be close to that of background;
- Many cluttered cells in a single view. A high number of occluding cells make image segmentation difficult;
- Low quality. Traditional staining techniques introduce a lot of inhomogeneity into the images, where not all of the parts of the same tissue or cells are equally stained.

Figure 1 shows a Papanicolaou stained cytological specimen indicating follicular carcinoma where in the characteristic problems are clearly evident. Image analysis approaches to cell segmentation are typically based on local image information such as regions, edges, histograms, or clusters [2, 3]. Edge-based or gradient-based segmentation methods rely on the idea of discontinuity of image intensities or texture at the boundary between different objects. Histogram-based and unsupervised clustering algorithms for image segmentation have been successful for those images with uniform illumination or other mapped discrimination features. Region-based algorithms employ region growing, region splitting and merging to separate different objects with homogeneous intensities. The edge-based methods and clustering approaches are generically sensitive to image noise and artifacts. The region-based algorithms, which are less sensitive to image noise, are usually computationally more expensive. These approaches are generally not applicable to cytological images because the grey level intensity of a cell image does not vary only on the boundary, but also within cells and throughout the background (see figure 1). Methods for segmentation that use region-based information are not applicable here because not all parts of the same tissue or cell cluster are equally stained. Darker background regions may be misclassified as cells and lighter cell regions may be misclassified as background. Edge-based segmentation methods are also not useful since boundary extraction does not provide good delineation of the cell boundary due to lack of contrast between the cell and background across all cells present in the image. Parametric techniques such as Hough transform based methods [4] have also not been successful for the same reasons.

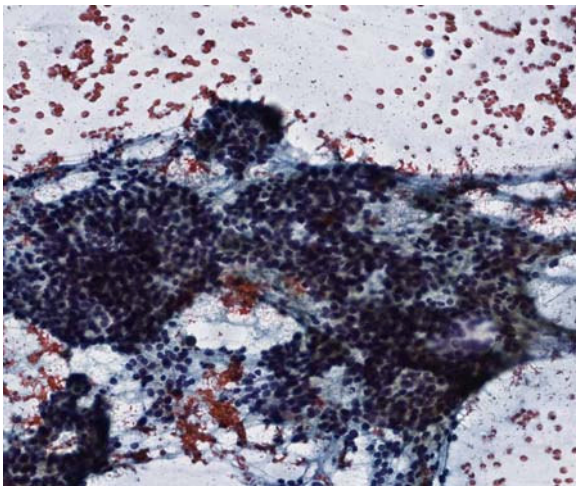

Figure 1: Image of a stained cytological smear.

Traditional image analysis methods have viewed segmentation as a low-level operation decoupled from higher level operations such as classification. However, the two processes are closely related. Each can be improved with information that the other provides. Such realization have been proposed in the domains of chromosome segmentation [5]. In this paper, we present a two-phase approach to segmenting cells in immunohistochemical stained cytolog-

ical images. We use an unsupervised clustering approach coupled with cluster merging based on a fitness function in the first phase to obtain a first approximation of the cells location. A joint segmentation-classification approach incorporating ellipse as a shape model is used in the second phase to detect the final cell contour. Specifically, the first phase formulation is based on the use of representative-based clustering coupled with cluster merging using proximity graphs. Second phase formulation is based on the Level Set approach proposed by Osher and Sethian [6] coupled with a feature-based classification model and the elliptical shape prior.

The remainder of the paper is organized as follows: section 2 describes the method to obtain a first approximation of the cell location. The algorithm yielding the final cell contour based on the level-set formulation is discussed in section 3. Section 4 outlines the experiments that were performed, and the results of segmentation. Finally, conclusions and a summary of our proposed approach appear in section 5.

2. Clustering and Cell Localization

The objective of the first phase analysis of cytological images is to find a set of locations corresponding to the cells of interest. We utilize the K-Means clustering algorithm that hierarchically splits the color space containing colors from the cells and background regions. One of the limitations of unsupervised approaches like K-Means is the inability to predict the true number of clusters as well as the lack of arbitrary cluster shape representation. We address this problem based on post-processing of the realized clusters, leading to a merge criteria to realize the final localization of cells.

Due to the variability in staining techniques for cytological images, the separation of cells and background is not apparent and surely not realized as distinct clusters. We utilize a post-processing technique that is similar to agglomerative hierarchical clustering in that we iteratively merge two candidates. However, it differs from a hierarchical clustering algorithm in that it merges neighboring clusters that enhance a given fitness function the most, and not necessarily merges clusters that are closest to each other. We define a fitness function that utilizes the principles of cohesion and separation. The fitness function used is

$$Q(x) = \texttt{Separation}(x)^{\delta} / \texttt{Cohesion}(x)^{(2-\delta)} \quad (1)$$

where, $\texttt{Separation}(x)$ is defined as the ratio of total inter-cluster distances across all clusters to the inter-cluster distance of the cluster of interest, and $\texttt{Cohesion}(x)$ is the ratio of the total intra-cluster distances across all clusters to the intra-cluster distance of the cluster of interest. The distances are measured as the L2-norm between a point in the cluster and the cluster center. In general, separation measures how well-separated a cluster is from other clusters while cohesion measures the tightness of a cluster. Hence, δ weighs the importance of distinct clusters to cluster compactness. The overall process of localization starts by realizing clusters using the K-Means clustering algorithm with a large k value. Next, using the cluster representatives, a proximity matrix is constructed using Gabriel-graphs. An entry is 1 in the proximity matrix, if two clusters are neighboring. Next, given all the possible merge candidates (indicated by an entry of 1 in the proximity matrix) the two maximizing the measure of fitness according to equation 1 are merged to form a new cluster. This process is repeated till only two clusters remain or the fitness measure stops improving. The final clustering results in the optimal separation of cells and background.

All pixels labeled as cells are separated into a new image and a blob coloring operation performed to count the total number of regions. Each region is than isolated as a new image for the next stage in segmentation.

3. Variational Segmentation Model

Variational methods have been developed as an approach to image segmentation that aim at minimizing the segmentation energy E represented by a real value. The segmentation energy generally measures how smooth the regions are, the similarity between the segmented image and the original one and the similarity between the obtained edges and the discontinuities of the original image. The basic idea is to start with initial boundary shapes represented in a form of contours given as $C(s) = \{(x(s), y(s)) : 0 \leq s \leq 1\}$, and iteratively modify them by applying shrink/expansion operations according to the constraints of the image. Those shrink/expansion operations, called contour evolution $\partial C/\partial t$, can be performed by the minimization of an energy function or by the simulation of a geometric partial differential equation (PDE). Level set approaches proposed by Osher and Sethian [6] move contours implicitly as a particular level, usually the zero level, of a function $\phi(x, y)$ defined on the spatial domain, such as $C \equiv \{(x, y) : \phi(x, y) = 0\}, \forall (x, y) \in \Omega$ where Ω denotes the entire domain of an image $I(x, y)$. Therefore, the evolution of a level set function $\partial\phi/\partial t$ represents the evolution of a set of contours $\partial C/\partial t$, and the contours C partition the image plane Ω into two subsets $\{\Omega_{in}, \Omega_{out}\}$ according to the sign of the level set function $\phi(x, y)$. Also, the defined contours can split or merge according to its topological changes as the level set function $\phi(x, y)$ grows or shrinks.

The Mumford-Shah model [7] has been regarded as a general model within variational segmentation methods. According to Mumford-Shah's conjecture, the image segmentation is a variational problem of finding an optimal piecewise-smooth approximation $f(x, y)$ of the given scalar image $I(x, y)$ and a set of boundaries C, such that the approximation $f(x, y)$ varies smoothly within the connected components of the subsets excluding the boundaries $\Omega \backslash C$. Chan and Vese proposed piecewise constant active contour model [8] based on Mumford-Shah segmentation model [7], given by

$$\begin{aligned} \frac{\partial \phi(x,y)}{\partial t} &= \delta_\epsilon(\phi(x,y))[\nu\kappa(\phi(x,y)) \\ &\quad -(I(x,y)-\mu_1)^2 + (I(x,y)-\mu_2)^2], \end{aligned} \quad (2)$$

where $\delta_\epsilon(.)$ denotes the regularized form of Dirac delta function, and $\{\mu_1, \mu_2\}$ denotes the mean of the image intensity I measured at the inside and the outside of the contours. Mumford and Shah proposed to solve the variational segmentation problem by minimizing the following global energy function:

$$E(f,C) \equiv \int_\Omega |I(x,y) - f(x,y)|^2 dxdy + \mu \int_{\Omega \backslash C} |\nabla f(x,y)|^2 dxdy + \nu |C| \quad (3)$$

The variational boundaries C have the role of approximating the edges of $I(x,y)$ by smoothing $f(x,y)$ only on $\Omega \backslash C$. The minimization of the global energy function approximates the image $I(x,y)$ with $f(x,y)$, smoothes $f(x,y)$, and reduces the length of boundaries $|C|$. The global energy function given in equation 3 for all regions can be generalized by:

$$E(f,C) \equiv \sum_i \int_{\Omega_i} e_i(x,y) dxdy + \mu \int_{\Omega_i \backslash C} |\nabla f_i(x,y)|^2 dxdy + \nu |C|, \quad (4)$$

where the objective function $e_i(x,y)$ determines the condition of region-based segmentation for each subset Ω_i for the piecewise-constant contours shown in equation 2.

3.1. Shape-Classification Model

Our goal here is to extend the energy functional 4 in order to force the level set to segment only the regions of interest, namely the cells constrained by specific image features and a known parametric shape. This is done in general by modifying the objective function $e_i(x,y)$ and adding a term E_{shape} to the global energy function that measures how well the level set represents the cell.

Let the (vector-valued) image intensity $\mathbf{I}$ be a multidimensional random variable given by $\mathbf{I} \in \Re^B$ where B denotes the dimension of the image intensity $\mathbf{I}$, which is also equivalent to the number of features extracted from the image $\mathbf{I}(x,y)$. We propose an objective function to measure how much an image pixel is likely to be an element of a subset using a probability density function (PDF) estimated from training samples. The objective function is given by

$$e_i(x,y) \equiv -\log(p_i(\mathbf{I}(x,y) + P(\Omega_i)), \forall (x,y) \in \Omega, \forall i, \quad (5)$$

where $p_i(\mathbf{I}) : \Re^B \rightarrow \Re$ denotes the multivariate conditional PDF of a vector-valued image intensity $\mathbf{I}$ on the condition that the image pixel $\mathbf{I}(x,y)$ is an element of the subset Ω_i, and $P(\Omega_i)$ denotes the a priori probability of the subset Ω_i. Based on the energy function shown in equation 4, minimizing the energy function E is equivalent to maximizing the a posteriori probability $P_i(\mathbf{I}(x,y))$ for each subset Ω_i.

Most human cells have elliptical shaped boundaries, but given the image quality, difficult to observe. As has been indicated previously, the ill-posed problem of this form can be solved by imposed parameter constraints in the form of *a priori* information. In a model-based approach, cell image segmentation can be cast as a parameter optimization problem. If the model parameters of a cell boundary are determined, we can reconstruct the segmented image, which is used to extract meaningful geometric parameters for pathologists. To impose such *a priori* knowledge, we use the implicit form of the ellipse equation to describe the cell boundary.The segmentation according to an ellipse then is equivalent to the recovery of five parameters, where the space to be optimized is given by $\Theta = [a, b, \theta, x_0, y_0]$ where (x_0, y_0) denotes the center of the ellipse, θ denotes the orientation of the ellipse, and a, b denote the major and minor axis of the ellipse, respectively.

Segmentation based on explicit incorporation of cell classification and its shape representation is now equivalent to deforming an ellipse according to Θ so it is attracted to the desired image classification. Within this formulation, the smoothness constraint is automatically ensured and therefore not needed from the original Mumford-Shah model. Hence, the modified global energy functional is given by:

$$E(f,C,\Theta) \equiv \sum_i \alpha \int_{\Omega_i} E_{class} + \beta \int_{\Omega_i} E_{shape} + \mu \int_{\Omega_i \backslash C} |\nabla f_i(x,y)|^2 dxdy, \quad (6)$$

where

$$E_{class} = \int_{\mathbf{I}} -\log(p_i(\mathbf{I}(x,y) + P(\Omega_i)) dxdy, \quad (7)$$

$$E_{shape} = \int_\Theta 1 - [(\frac{[(x-x_0)\cos\theta + (y-y_0)\sin\theta]^2}{a^2} + \frac{[(x-x_0)\sin\theta + (y-y_0)\cos\theta]^2}{b^2})]^{\frac{1}{2}} dxdy, \quad (8)$$

and α and β weight the contribution of the classification and shape information in the energy functional.

3.2. Curve Evolution

By setting the contour pixels as an element of interior subset, such as $\Omega \equiv \bigcup_i \Omega_i$, each subset on the spatial domain can be identified by a set of binary identity functions

$$\chi_i(x,y) \equiv \left\{ \begin{array}{c} 1, \text{if}(x,y) \in \Omega_i \\ 0, \text{otherwise} \end{array} \right\}, \forall i, \quad (9)$$

composed of a group of regularized unit step functions $\{H_j\}$ given by

$$H_j \equiv H_\epsilon(\phi(x,y)) \approx \left\{ \begin{array}{ll} 1, & \text{if} \phi_j(x,y) \geq 0 \\ 0, & \text{if} \phi_j(x,y) < 0 \end{array} \right\}, \quad (10)$$

for $\forall (x,y) \in \Omega, \forall j$. Using the identity function, the integration over each subset Ω_i is generalized to the integration over the entire image plane Ω. Also, the shape of contours C_j is equivalent to the integration of E_{shape} over the image plane where C_j denotes a set of active contours formed by

the corresponding level set function $\phi_j(x,y)$. The global energy function of the level set contour model and the associated Euler-Lagrange equation obtained by minimizing the energy function E with respect to $\phi = \phi_1, \ldots, \phi_j, \ldots, \phi_J$ in [8] can be extended for our proposed global energy function as

$$E = \sum_{i=0}^{m-1} \alpha \int_{\Omega} E_{class}\chi_i(x,y)dxdy + \beta \int_{\Omega_i} E_{shape}\chi_i(x,y)dxdy + \mu \int_{\Omega} f_i(x,y)\chi_i(x,y)dxdy, \quad (11)$$

and

$$\frac{\partial \phi_j(x,y)}{\partial t} = \delta_j[-\mu - \sum_{i=0}^{m-1} \alpha E_{class}\frac{\partial \chi_i}{\partial H_j} - \sum_{i=0}^{m-1} \beta E_{shape}\frac{\partial \chi_i}{\partial H_j}], \forall j, \quad (12)$$

where $\delta_j \equiv \delta_\epsilon(\phi_j(x,y))$. The proposed contour evolution model is obtained by solving for the relevant objective functions E_{class} and E_{shape} shown in equations 7 and 8, respectively, into the contour model shown in equation 12, given by

$$\frac{\partial \phi_j(x,y)}{\partial t} = \delta_\epsilon(\phi(x,y))[-\mu - \alpha(\sum_{i=0}^{m-1} P_{i,cell}(\mathbf{I}(x,y)) - \sum_{i=o}^{m-1} P_{i,bckg}(\mathbf{I}(x,y))) - \beta(\sum_{i=0}^{m-1}(1 - \sqrt{(A/a)^2 + (B/b)^2}))] \quad (13)$$

where, $A = (x_i - x_0)\cos\theta + (y_i - y_0)\sin\theta$ and $B = -(x_i - x_0)\sin\theta + (y_i - y_0)\cos\theta$. The solution to the Euler-Lagrange is implemented using gradient descent where the parameters for the ellipse, Θ, are solved at each iteration of the level set evolution, given as:

$$\begin{aligned}
\frac{\partial a}{\partial t} &= -\int_{\Omega} \delta_\epsilon(\phi_j(x,y))A(1/a^3)dxdy \quad (14)\\
\frac{\partial b}{\partial t} &= -\int_{\Omega} \delta_\epsilon(\phi_j(x,y))B^2(1/b^3)dxdy\\
\frac{\partial x_0}{\partial t} &= -\int_{\Omega} \delta_\epsilon(\phi_j(x,y))(A\cos\theta/a^2) - B\sin\theta/b^2)dxdy\\
\frac{\partial y_0}{\partial t} &= -\int_{\Omega} \delta_\epsilon(\phi_j(x,y))(A\sin\theta/a^2) + B\cos\theta/b^2)dxdy\\
\frac{\partial \theta}{\partial t} &= -\int_{\Omega} \delta_\epsilon(\phi_j(x,y))(AB[1/b^2 - 1/a^2])dxdy.
\end{aligned}$$

3.3. Image Features and Pixel Classification

The objective of the energy function E_{class} is to partition the image such that the local image statistics within the cell and background are "close" to the global statistics within the same class across the image. This is expressed according to Bayesian decision rule formulated as the *maximum a posteriori* solution. In deriving the Bayesian classifier, low-level features of the image are computed as characteristics in defining the likelihood function. Specifically, color and texture features are computed from the color image and the converted gray-level image, respectively. For the color features, we use 1976 CIE $L^*a^*b^*$ color space separated into luminance and chrominance channels. Color distribution is modeled with histograms constructed with kernel density estimates. Histograms are compared with the χ^2 histogram difference operator [9] to obtain a feature for each pixel. A brightness cue is also computed based on the use of L^* histogram for each pixel. Once again, the actual feature for each pixel is the χ^2 histogram difference. Texture features used are computed from the gray-level image based on gray-level co-occurrence matrices, fractal measures, Law's texture measures, gradient structure tensors, and Gabor filters [10]. The likelihood functions for the cell region and background are modeled using a mixture of Gaussians [10] and the priors computed as a ratio of the average number of cell pixels to background pixels in the training image set.

4. Results

The proposed method has been evaluated to segment immunohistochemically stained cytological smears for thyroid lesions. A total of 50 images were available, each manually segmented to delineate the cells of interest. 10 images with sufficient variability in background and stain conditions were used to compute low-level image features and generate the likelihood models and estimates of prior probabilities to be used in cell-background classification (section 3.3). In addition, the specific stain color used to localize cells of interest was also noted to select the appropriate cluster as a result of the localization step described in section 2.

The learned parameters were used to automatically segment the remaining 40 images. Each localized region was isolated as a new image to be segmented. The center of the isolated region and $1/4$ of the region height and width was used to initialize the level set contour. The contour was evolved according to the equations in section 3.2 to obtain the final segmentation where each region was partitioned into two subsets $\Omega_{in}, \Omega_{out}$ based on the zero level set. Figure 4 shows three representative localized regions and the corresponding cell segmentation obtained based on the proposed level set model. Images used for test resulted in cell segmentation accuracy rate of 92.1% with a missed segmentation rate of 4.3% and a false segmentation rate of 2.7%. The segmentation errors encountered were primarily due to the failure of the localization stage in identifying the correct region of interest.

5. Summary and Conclusions

We have proposed an advanced segmentation method using curve evolution based on level set theory combined with a Bayesian classification model and *a priori* shape knowledge. A two-phase approach to segmenting cells

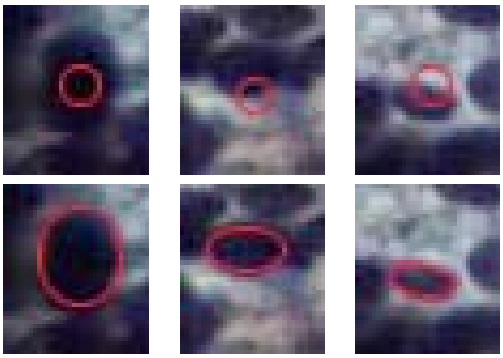

Figure 2: Segmentation of three cells within the corresponding localized regions. The top row shows the localized regions and the contour initialization. The bottom row shows the final contour for each of the regions.

in immunohistochemical stained cytological images is presented. An unsupervised clustering approach coupled with cluster merging based on a fitness function is used as the first phase to obtain a first approximation of the cells location. A joint segmentation-classification approach incorporating ellipse as a shape model is used as the second phase to detect the final cell contour. For pixel-classification, the segmentation model estimates a multivariate density function of low-level image features from training samples and uses it as a measure of how likely each image pixel is to be a cell. This estimate is constrained by the zero level set, which is obtained as a solution to an implicit representation of an ellipse. The developed method is tested on 40 cytological images of thyroid lesions and results compared to manual delineation of cells.

References

[1] H.S. Wu, J. Barba, and J. Gil, "A parametric fitting algorithm for segmentation of cell images," *IEEE Trans Biomed Eng*, vol. 45, pp. 400–407, 1998.

[2] R. Adams and L. Bischof, "Seeded region growing," *IEEE Trans. Pattern Analys. Machine Intell.*, vol. 16, pp. 641–647, June 1994.

[3] W.N. Lie, "Automatic target segmentation by locally adaptive image thresholding," *IEEE Trans. Image Process.*, vol. 4, pp. 1036–1041, July 1995.

[4] P.K. Ser and W.C. Siu, "Novel detection of conics using 2-D Hough planes," in *IEEE Proc. Vision, Image Signal Processing*, 1995, vol. 142, pp. 262–270.

[5] G. Agam and I. Dinstein, "Geometric separation of partially overlapping nonrigid objects applied to automatic chromosome classification," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 11, pp. 1212–1222, November 1997.

[6] S. Osher and J. Sethian, "Fronts propagating with curvature dependent speed: Algorithms based on Hamilton-Jacobi formulations," *Journal of Computationl Physics*, pp. 12–49, 1988.

[7] D. Mumford and J. Shah, "Optimal approximation by piecewise smooth functions and associated variational problems," *Communication Pure and Applied Mathematics*, p. 577, 1989.

[8] L. Vese and T. Chan, "A multiphase level set framework for image segmentation using the Mumford and Shah model," *International Journal of Computer Vision*, , no. 3, pp. 271–293, 2002.

[9] J. Puzicha, T. Hofmann, and J. M. Buhmann, "Histogram clustering for unsupervised segmentation and image retrieval," *Pattern Recognition Letters*, vol. 20, no. 8, pp. 899–909, Aug. 1999.

[10] S. Shah and J. K. Aggarwal, "Multiple feature integration for robust object localization.," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 1998, pp. 765–771.