Automatic Cell Segmentation Using a Shape-Classification Model in Immunohistochemically Stained Cytological Images

Shishir SHAH\textsuperscript{a)}, Nonmember

**SUMMARY** 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. 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. Results of segmentation are presented and compared to ground truth measurements.

**Key words:** image segmentation, probabilistic curve evolution, image cytology

1. Introduction

Evaluation of cytology specimen for the purpose of making diagnostic decisions is guided by the measurement of cell properties and their spatial distribution. In particular, the geometric properties of cells such as area, radius, perimeter, convexity, etc., and their spatial distribution measured in terms of the specimen topology play an important role in the diagnostic process [1]. Thus, in developing a computer-assisted system for cytology specimen analysis, segmentation and delineation of cells forms the first step necessary for accurate quantification of cell parameters.

Automatic segmentation of cells in a multitude of image modalities has been a problem of interest over the last three decades [2]. While many approaches have been proposed [1], [3], including specific methods for images of immuno-stained cytology specimen [4], cell segmentation remains a problem of interest due to both the complex nature of cell structures with their associated variabilities, and the problems inherent in the imaging process. Current practice for evaluation of a cytological specimen requires that a cell smear be appropriately prepared and stained. The stains used label (modify the visible appearance of) the cells and tissue fragments of interest. Following staining, the specimen are digitally imaged using a light microscope. The next step involves analysis of the digitized image(s) to obtain relevant measurements leading to a diagnostic outcome. Common problems encountered in immunohistochemical stained cytology specimen images include object multiplicity, short range of grey levels, clutter, occlusion, and non-random noise. These are characterized by:

- Poor contrast, i.e., cell grey levels generally bleed into neighboring or overlapping cells making the observation of a distinct cell boundary difficult;
- Many cluttered cells in a single view. Cells tend to group together unless specifically prepared for isolation. This results in overlapping cells with partial view or complete occlusions;
- 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. Moreover, this problem is exemplified due to inconsistent staining from one specimen preparation to the next.

Figure 1 shows a Papanicolaou stained cytological specimen indicating follicular carcinoma where in the characteristic problems are clearly evident.

One of the most common approaches to cell segmentation is based on thresholding [5]. While segmentation based on such an approach is used often for cytological images, it is semi-automated at best and requires frequent readjustment of threshold values due to the inherent variability in the staining process that results in different grey level distributions for different images. Further, the grey level intensity varies even within a cell making it difficult to delineate an entire cell based on a single threshold value (see Fig. 1).

Approaches that leverage local image information such as regions, edges, histogram, and clusters have also been developed for the purpose of cell segmentation [6]. Methods that are edge- or gradient-based rely to the notion that dis-

---

\textsuperscript{a}) E-mail: shah@cs.uh.edu

Manuscript received November 2, 2007.

\textsuperscript{1}The author is with the Department of Computer Science, University of Houston, Houston, TX, U.S.A.

DO: 10.1093/ietisy/e91-d.7.1955

Copyright © 2008 The Institute of Electronics, Information and Communication Engineers
continuities in the image signify boundaries between different objects. Objects that exhibit uniform intensity values or a homogenous distribution of intensities have been successfully segmented using histogram- or clustering-based approaches. All of these methods are known to be sensitive to noise and other image artifacts. More specifically, edge-based and histogram-based approaches are not useful for cytological smears 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. Modified approaches that combine edge-based methods with parametric techniques such as Hough transform [7] have also not been successful for the same reasons. Image textures and noise usually degrade edge detection making it difficult to find the correct peaks in the parameter space. On the other hand, region-based approaches that employ region growing, splitting, and merging algorithms, are less sensitive to noise. However, due to their iterative nature, they tend to be computationally more expensive. Furthermore, cytological smears typically contain cell clusters that are not equally stained making region-based approaches difficult to adapt resulting in darker background regions being classified as cells and lighter cell regions being misclassified as background.

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. An example of such coupled interaction between low- and high-level processes has been proposed in the domain of chromosome segmentation [8]. In this paper, we present a two-phase approach to segmenting cells in immunohistochemically stained 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 boundary. 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 [9] coupled with a feature-based classification model and the elliptical shape prior.

The remainder of the paper is organized as follows: Sect. 2 describes the method to obtain a first approximation of the cell location. The algorithm yielding the final cell boundary based on the level-set formulation is discussed in Sect. 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 Sect. 5.

2. Clustering and Cell Localization

The objective of the first phase analysis is to find a set of locations corresponding to the cells of interest. Due to problems in the images (noise, overlapping cells, etc.), a natural way of doing this is to use representative techniques that account for pixel-level information. 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 group clusters into arbitrary shapes not necessarily dependent only on the spatial locations of points in the individual clusters. This leads to a coherent grouping of points defining a separation of cells from the image background.

Due to the variability in staining techniques for cytological images, the separation of cells and background is not realized as distinct clusters. Similar to agglomerative hierarchical clustering approaches, the initial clusters are post-processed to iteratively merge two candidate clusters. However, the approach used differs from a traditional hierarchical clustering algorithm in that two clusters are merged only if a given fitness function is maximized, and not necessarily merged due to their proximity to each other. Specifically, the fitness function used is one that utilizes the principles of cohesion and separation [10]. The fitness function is given by:

\[ Q(x) = \frac{\text{Separation}(x)^\delta}{\text{Cohesion}(x)^{(2-\delta)}} \]  \hspace{1cm} (1)

where, Separation(x) is defined as the ratio of total inter-cluster distances across all clusters to the inter-cluster distance of the cluster being considered, and Cohesion(x) is the ratio of the total intra-cluster distances of all clusters to the intra-cluster distance of the cluster being considered. In computing the inter- and intra-cluster distances, the distance metric used is the L2-norm and all distances are measured between a point in the cluster and the cluster center. Separation provides a measure of how different a cluster is from rest of the clusters while cohesion provides a measure of how similar are the points belonging to a cluster. The weight factor \( \delta \) indicates the importance of having distinct clusters with a homogenous set of points. Typically, \( \delta \) takes a value that is greater than or equal to 1 but less than 2.

The overall process of localization starts by partitioning the image into clusters using the K-Means clustering algorithm that is initialized with a large k value. Next, a proximity matrix is constructed for all the representative clusters using Gabriel-graphs. If two clusters are neighbors, the proximity matrix is updated to reflect a value of 1. This is repeated for all clusters. A value of 1 in the proximity matrix indicates all possible merge candidates. Next, for all entries of 1 in the proximity matrix, an exhaustive search is performed to identify and merge the two clusters that maximize the fitness according to Eq. (1). The proximity matrix is once again updated based on the remaining clusters. This process is iterated till the fitness measure stops growing.

Of the resulting clusters, the cell cluster is selected based on a priori knowledge of typical cell characteristics,
such as the stain color. Pixels belonging to the selected cluster are separated to form a new image of the same size as the original image. A watershed region partitioning is performed followed by edge-constrained region growing [11], [12]. Finally a blob coloring operation is performed to count the total number of regions [13]. Each region is than isolated as a new image for the next stage in segmentation.

3. Variational Segmentation Model

Among traditional image segmentation techniques, active contour methods have been developed that transform the region-partitioning objective to one of minimizing an energy function $E$ represented by a real value. The segmentation energy generally measures how smooth the segmented regions are and the similarity between the obtained edges and the discontinuities of the original image. The process is initialized with the definition of an initial contour that describes the boundary of the region to be segmented. The contour can be given as:

$$C(s) = \{(x(s), y(s)) : 0 \leq s \leq 1\}. \quad (2)$$

The contour is evolved iteratively by applying a set of operations that allow the contour to move inwards or outwards based in the imposed image constraints. Two approaches are most prevalent in evolving the initial contour. One is based on the classic snakes formulation proposed by Kass et al. [14] that moves the contour points to minimize an energy functional. The other is the level set approach proposed by Osher and Sethian [9] that moves the contour implicitly based on the zero-level of the function defined on the spatial domain. Level set formulations are advantageous over classic snakes when initial contour position cannot be accurately defined and when multiple regions need to be segmented simultaneously. Several segmentation methods have been developed based on contour evolution [15], [16].

3.1 Segmentation Model

The Mumford-Shah model [17] has been regarded as a general model within variational segmentation methods. Accordingly, the segmentation problem is formulated as one 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. Based on this definition, Mumford and Shah defined the following energy function to be minimized that would solve the defined segmentation problem:

$$E(f, C) = \int_{\Omega} |I(x, y) - f(x, y)|^2 dx\,dy + \mu \int_{\Omega \cap C} |\nabla f(x, y)|^2 dx\,dy + \nu |C| \quad (3)$$

The first integral defines the best function that describes the image intensity within the subset while the second integral ensures smoothness of the function to ensure smooth boundaries. Overall, the contour $C$ found by minimizing the above energy function will approximate the edges of $I(x, y)$ by smoothing $f(x, y)$. The second term in this general formulation can be ignored in many cases since the third term also defines smoothness based on the contour length. Without loss of generality, the global energy function given in Eq. (3) can be re-written for any arbitrary objective function $e_i(x, y)$ that defines the region-based segmentation over all regions $\Omega_i$ as:

$$E(f, C) = \sum_i \int_{\Omega_i} e_i(x, y) dx\,dy + \mu \int_{\Omega \cap C} |\nabla f_i(x, y)|^2 dx\,dy + \nu |C|, \quad (4)$$

3.2 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 $E_{\text{shape}}$ to the global energy function that measures how well the level set represents the cell.

Since the objective function is meant to measure how much the pixels within the boundary $C$ belong to a defined function, we can generalize the objective function to measure different properties of the pixel besides its intensity. Let us consider a pixel’s attribute to be a multidimensional random variable represented by $\vec{I}$. The actual attribute for a pixel can include any feature extracted from the image. Using Bayesian decision theory, the objective function becomes the likelihood of a pixel belonging to the probability density function (pdf) of the cell. This pdf would be derived from training samples comprised of attributes of the pixels that belong to the cell region. The objective function can then be given as:

$$e_i(x, y) = -\log(p_i(\vec{I}(x, y) + P(\Omega_i)), \quad (5)$$

where $p_i(\vec{I})$ is the multivariate pdf conditioned such that the pixels considered are within the domain $\Omega_i$ and $P(\Omega_i)$ is the prior probability of observing the domain in the image. Replacing this as the objective function in Eq. (4), minimizing the global energy function $E$ is equivalent to maximizing the a posteriori probability $P_i(\vec{I}(x, y) + \Omega_i)$ for each subset $\Omega_i$.

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 imposing parameter constraints in the form of a priori information. One shape constraint that can be easily embedded within the boundary contour defined according to the global energy function is given by the implicit equation of an ellipse. This can be written as:

$$\frac{(x - x_0)\cos \theta + (y - y_0)\sin \theta}{a^2} = 1,$$
where \((x_0, y_0)\) denotes the center of the ellipse, \(\theta\) denotes the orientation of the ellipse, and \(a, b\) denote the major and minor axis of the ellipse, respectively. For a given set of contour points, the bounding ellipse can be obtained by recovering the five unknown parameters noted by \(\Theta = [a, b, \theta, x_0, y_0]\).

Segmentation based on explicit incorporation of cell classification and its shape representation is now equivalent to deforming an ellipse according to \(\Theta\) so it is attracted to the boundary points surrounding the image region obtained according to pixels classified by the objective function. 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:

\[
E(f, C, \Theta) = \sum_i \alpha \int_{Q_i} E_{\text{class}} + \beta \int_{Q_i} E_{\text{shape}} + \mu \int_{Q_i \setminus C} |\nabla f_i(x, y)|^2 dxy,
\]

where

\[
E_{\text{class}} = \int_f -\log(p(t(x, y) + P(\Omega_i)) dxy,
\]

\[
E_{\text{shape}} = \int_\Theta \left[ \left( \frac{(x-x_0)\cos(\theta)}{a^2} + 
\frac{[x-x_0)\sin(\theta) + (y-y_0)\cos(\theta)]^2}{b^2} \right)^{\frac{1}{2}} \right] dxy,
\]

and \(\alpha\) and \(\beta\) weight the contribution of the classification and shape information in the energy functional.

### 3.3 Bayesian Classification

In solving for the posterior probability required by the objective function \(e(x, y)\), we discuss the design of a single classifier for the two-class discrimination case, where the first class is the cell and the second class is background. Since multiple attributes are computed for each pixel, let us consider the transformed feature space as \(Q\), where each attribute \(q\) is sampled from the space of features \(S\) and is a statistical property of the image intensity. As we are interested in modeling the class signatures, a set of training features are used that are extracted from known classes in the image. Thus the distribution of the features is given as the class conditional density function \(p(q|t)\), where \(t\) comes from the true class space \(T\), with a priori distribution \(p(t)\).

In the Bayesian framework, the decisions are made by evaluating the \textit{a posteriori} probability for each class and choosing the one with the highest probability as the true class. Given the \textit{a priori} probability of any class pixel \(P(t)\) and the conditional density \(p(q|t)\), we can compute the posterior probability of the observed feature being a particular class pixel using the Bayes rule. For the two-class discrimination case, the posterior probability given by the classifier is:

\[
P(t|q) = \frac{p(q|t)P(t)}{p(q|t)P(t) + p(q|b)P(b)}
\]

where, \(p(q|b)\) is the conditional density function of the background distribution and \(P(b)\) is the prior probability of observing the background. Assuming a parametric model for the class conditional pdf’s, the only remaining unknown parameter is the prior probability of observing a class region. This is calculated from a training set of images. As we are performing detection based on individual pixels, the prior probability is computed by:

\[
P(t) = \frac{\# \text{ of cell pixels}}{\text{total \# of image pixels}}
\]

and

\[
P(b) = 1 - P(t)
\]

To achieve optimum performance from any classification/clustering system, it is essential that its design exploits the specific characteristics of the data. Since multiple attributes are computed for each pixel, the class conditional pdff has to appropriately model the distribution of features in a high-dimensional space. If each class dependent distribution is derived from a homogeneous image intensity distribution, a unimodal probability distribution such as a multivariate Gaussian distribution may be good enough as \(p(I)\) [16]. However, the statistical property of each class conditional distribution is often nonuniform due to the complexity of the cells and background themselves and the presence of noise in the data. Hence feature reduction is necessary to avoid the drawbacks of Occam’s razor [18]. Instead of designing a single classifier that accounts for all the pixel attributes within a single pdf, we propose a scheme that uses individual classifiers for each subset of the attributes and fuses the outputs of all classifiers to reach a consensus. Each individual classifier provides an estimate of the posterior probability and the goal of the combining stage is to produce a single estimate that maximizes the probability for localized cell detection while reducing clutter and false alarms. Various integration methods have been proposed in the past [19]. We formulate a supra-Bayesian integration in which the posterior estimates from each classifier are assumed to have a probability distribution and, based on the means and variances of the outputs, we can formulate an optimal decision scheme. Strictly speaking, Bayesian theory holds true only for individual decision makers, but if the group decision is viewed as a collaborative effort, the effect is externally Bayesian. As in the case of individual classifiers, the integration module is estimating the probability of observing a cell or background pixel. So, for \(n\) individual classifiers, where each \(P(t|q)\) is providing a measure of subjective probability of observing a cell pixel, and if the
posterior are Gaussian distributed, the integrated posterior decision simplifies to:

$$P(t|q_1, q_2, \ldots, q_n) = \frac{\prod_{i=1}^{n} P(q_i|t)}{\prod_{b=1}^{n} P(q_i|b)} P(t) + \frac{\prod_{i=1}^{n} P(q_i|b)}{\prod_{b=1}^{n} P(b)} P(b)$$

(13)

where, $w_i$ weights the contribution of each of the features.

As each of the classifiers is designed to identify single class pixels, we know that there is sufficient diversity and complementarity within the estimates. Thus the weight associated with each of the classifiers plays an important role in deciding the contribution from each estimate. This is mainly due to the fact that the integrator module does not have the same information that is seen by each of the classifiers. Evaluating the log likelihood of Eq. (13) and assuming that the combined probability ratios provide the final probability as

$$P \left( \ln \left( \frac{P(t|q_1)}{1 - P(t|q_1)} \right), \ldots, \ln \left( \frac{P(t|q_n)}{1 - P(t|q_n)} \right) \right)$$

(14)

and

$$P \left( \ln \left( \frac{P(b|q_1)}{1 - P(b|q_1)} \right), \ldots, \ln \left( \frac{P(b|q_n)}{1 - P(b|q_n)} \right) \right)$$

(15)

and, if the joint distributions are multivariate normal densities with mean $\mu_t$ and $\mu_b$ and covariance $\Sigma_t$, then the weights for the individual classifiers can be computed by:

$$w_i = \Sigma_t^{-1}(\mu_t - \mu_b)$$

(16)

This result provides an intuitive insight to the integration of decisions. In general, when all the classifiers provide similar estimates, the combining results in the peaking of that estimate. On the other hand, and more importantly, when the classifiers do not agree on an estimate, their reliability has to be considered. According to the weight assignment in Eq. (16), the reliability associated with each of the classifiers will depend on how different its estimate is from rest of the classifiers, and how much diversity exists within the estimates.

### 3.4 Mixture Density Model

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 grey-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 $\chi^2$ histogram difference operator [20] 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 $\chi^2$ histogram difference. Uniformity, entropy, contrast, inverse differential moment, and correlation are computed based on grey-level co-occurence matrices [21]. Features are also derived from the fractal dimension analysis of a texture [22]. Finally, we compute two additional measures of texture based on Laws’ features and Gabor filters [23]. Due to the complex and non-Gaussian distribution of each of the five feature vectors, we model the data using a mixture of Gaussians.

Modeling of data is an important consideration in designing statistical classifiers. The simplest way to model non-Gaussian data is to use the histograms of the training data. However, classification based on this method does not generalize well from the training data to the test data. The Parzen density estimate is a well established method to establish density estimates for multivariate models. However, the Parzen windows approach is computationally expensive and has problems when the data is large and sparsely distributed. We use the Expectation-Maximization (EM) algorithm [24] to determine the parameters for the mixture of Gaussians model to estimate the density function. Considering $Q$ to be the data, we pose the parameter estimation as a maximum likelihood problem. The general form of the density function for the measured feature can be given as:

$$P(Q|t) = \sum_{i=1}^{c} p(Q|t, \theta_i) \alpha_i$$

(17)

where, $t$ is the conditioning variable (class signature), $c$ represents the number of component density functions $p(Q|t, \theta_i)$ that make up the mixture, $\alpha_i$ represents the weight associated with each of the density functions (also called mixing parameter), and $\theta_i$ represents the parameter vectors for each component density function. $\theta$, $\alpha$, and $c$ are unknown, and have to be estimated from the data. We assume the component densities to be normal distributed. That is $p(Q|t, \theta_i) \approx N(\mu_i, \Sigma_i)$, and $\theta_i = (\mu_i, \Sigma_i)$, where $\mu_i$ and $\Sigma_i$ represent the multivariate mean and covariance matrix of the normal distribution. We use the K-Means algorithm iteratively with the EM algorithm to determine all the parameters [11].

### 3.5 Curve Evolution

Chan and Vese have already shown the solution to the global energy function (Eq. (3)) of the level set contour and the associated Euler-Lagrange equations obtained by minimizing the energy function in [15]. Without loss of generality, the proposed solution can be extended to curve evolution for any arbitrary objective function $e(x, y)$. Extending this further, the modified shape-classification global energy function of the level set contour model and the associated Euler-Lagrange equation can be given by:

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

(18)
and

$$\frac{\partial \phi_j(x,y)}{\partial t} = \delta_j \left[ -\mu - \sum_{j=0}^{m-1} \alpha E_{class} \frac{\partial \chi_j}{\partial H_j} \right. \\
\left. - \sum_{j=0}^{m-1} \beta E_{shape} \frac{\partial \chi_j}{\partial H_j} \right], \forall j, \tag{19}$$

where \(\delta_j \equiv \delta_k(\phi_j(x,y))\), wherein \(\delta_k(.)\) denotes the regularized form of Dirac delta function and \(\phi(x,y)\) is the level set function, \(\chi_j\) is a binary identity function that is 1 if the pixel is within the domain \(\Omega_j\) and 0 otherwise, and \(H_j\) is a regularized unit step function that is 1 is outside the contour and 0 inside.

The final contour evolution model is obtained by solving for the relevant objective functions \(E_{class}\) and \(E_{shape}\) shown in Eqs. (8) and (9), respectively, into the contour model shown in Eq. (19), given by

$$\frac{\partial \phi_j(x,y)}{\partial t} = \delta_k(\phi_j(x,y))$$

$$\cdot \left[ -\mu - \alpha \sum_{i=0}^{m-1} P_{i,cell}(\tilde{H}(x,y)) - \sum_{i=0}^{m-1} P_{i,bckg}(\tilde{H}(x,y)) \right]$$

$$- \beta \left( \sum_{i=0}^{m-1} \left( 1 - \sqrt{(A/a)^2 + (B/b)^2} \right) \right), \tag{20}$$

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, \(\Theta\), are solved at each iteration of the level set evolution [25], given as:

$$\frac{\partial a}{\partial t} = -\int_{\Omega} \delta_k(\phi_j(x,y)) A(1/a^2) dxdy$$

$$\frac{\partial b}{\partial t} = -\int_{\Omega} \delta_k(\phi_j(x,y)) B^2(1/b^2) dxdy$$

$$\frac{\partial x_0}{\partial t} = -\int_{\Omega} \delta_k(\phi_j(x,y)) (A \cos \theta/a^2) - B \sin \theta/b^2) dxdy$$

$$\frac{\partial y_0}{\partial t} = -\int_{\Omega} \delta_k(\phi_j(x,y)) (A \sin \theta/a^2) + B \cos \theta/b^2) dxdy$$

$$\frac{\partial \theta}{\partial t} = -\int_{\Omega} \delta_k(\phi_j(x,y)) (AB(1/b^2 - 1/a^2)) dxdy. \tag{21}$$

4. Results

The proposed method has been evaluated on a image set of 50 immunohistochemically stained samples, of which 18 samples were stained with H&E, 15 with Papanicolaou, and 17 with DAB chromogen. Each of the images were manually segmented to outline individual cells. The image set was divided into a training set of 10 images and a test set of 40 images, such that each stain type was represented in each set. The training set was used to estimate parameters for both segmentation phases, learn the color for cell clusters to differentiate from the background cluster at the end of the localization step (Sect. 2), generate the likelihood models, and estimate prior probabilities to be used in cell-background classification step (Sect. 3.3). The parameters for the two phases were manually adjusted to optimize the segmentation accuracy on the training set. For the localization stage, the initial \(k\) value was set to 1,000, and the \(\delta\) value set to 1.2 for the fitness function. Similarly, the weight factors for the the classification and shape energy components of the shape-classification energy functional were determined to be \(\alpha = 0.38\) and \(\beta = 0.62\), respectively.

The same parameters were used to segment images in the test set. Figure 2 shows the result of localization and sample cell segmentations on the image in Fig. 1. The image (a) shows the pixels within the identified cell cluster and the partitioning obtained at the end of the localization phase. Each region delineated by the white boundary is identified as a potential region containing a cell and processed individually by the next step for further segmentation. A bounding rectangle containing the localized region is extracted from the original image and the level set contour initialized at the center of the region, having a diameter that is equal to the greater of 1/4 of the region height or width. The contour was evolved according to the equations in Sect. 3.5 to obtain the final segmentation where each region was partitioned into two subsets \(\Omega_{in}, \Omega_{out}\) based on the zero level set. Example of three cell regions with overlayed initial and final contour are shown in columns (b) and (c) of Fig. 2.

Figure 3 (a) shows another example image with its corresponding segmentation shown in Fig. 3 (b). This example illustrates the limitations of the proposed approach. The image in this example exhibits contrast problems due to poor staining quality of the specimen. In addition, the cells themselves show fragmented nucleic structures and hence deviate from our assumption of the elliptical shape model. The first cell regions shown in column (2) of Fig. 3 (b) is identified correctly by the localization stage (Fig. 3 (b)(1)), but the segmentation (shown in column (3)) fails to identify the correct cell boundary because the poor staining quality results in significant number of cell pixels having properties similar to that of background pixels, coupled with a weak cell edge. On the other hand, regions shown in the second and third row of Fig. 3 (b)(2) are incorrect localizations. The first re-
region contains two cells that are very close to each other and due to incorrect localization, only one cell is segmented. For the second region, the localization stage identifies a region that does not contain a cell, but due to bleeding of the stain, the background region has pixels with the same color as that of the cell pixels. This results in a false detection of a cell.

Overall, for the images in the test set, the algorithm resulted in a cell segmentation accuracy rate of 92.1% with a false segmentation rate of 2.7%. Each cell segmented was examined to tabulate the error rates and to identify the reason for the error. We found that the primary reason for segmentation errors was failure of the localization stage in identifying the correct region of interest. This indicates that color alone may not provide sufficient differentiation of cell pixels, especially considering that the staining process is highly variable. In addition, it is also clear that a strict adherence to an elliptical shape model biases the ability to segment cells that deviate significantly from the normal structure.

Finally, to compare the proposed approach to existing segmentation methods, we implemented and tested the parametric fitting approach proposed by Wu et al. [1] on the 40 test images. Figure 4 shows the segmentation obtained for the image in Fig. 1. Due to the inability of the algorithm to perform localized analysis of finer cell structures. Overall, for the 40 images analyzed, the segmentation accuracy obtained was 78.7% with a false segmentation rate of 11.3%. This clearly indicates the advantage of a two-stage segmentation process and the benefit of a combined shape-classification model that can perform local analysis.

5. 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 in immunohistochemically 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

1962

Shishir Shah received his Ph.D. degree in Electrical and Computer Engineering in 1998 from The University of Texas at Austin. He joined the Department of Computer Science at University of Houston in 2005, where he is an Assistant Professor. His research includes computer vision and pattern recognition with applications in the biomedicine and distributed multimodality sensing. He has co-edited one book, and authored papers on object recognition, sensor fusion, statistical pattern analysis, and bioinformatics.