Journal of Imaging Science and Technology® 52(4): 040502–040502-11, 2008.

Segmentation of Brain Immunohistochemistry Images Using Clustering of Linear Centroids and Regional Shapes

Hai-Shan Wu, Jacinta Murray and Susan Morgello
Department of Pathology, Box 1194, Mount Sinai School of Medicine, One Gustave L. Levy Place, New York, New York 10029
E-mail: haishan.wu@mssm.edu

Abstract. *A generalized clustering algorithm utilizing the geometrical shapes of clusters for segmentation of colored brain immunohistological images is presented. To simplify the computation, the dimension of vectors composed from the pixel RGB components is reduced from three to two by applying a de-correlation mapping with the orthogonal bases of the eigenvectors of the auto-covariance matrix. Since the brain immunohistochemical images have stretched clusters that appear long and narrow in geometrical shape, we use centroids of straight lines instead of single points to approximate the clusters. An iterative algorithm is developed to optimize the linear centroids by minimizing the approximation mean-squared error. The partitioning of the two-dimensional vector domain into three portions classifies each image pixel into one of the three classes: The microglial cell cytoplasm, the combined hematoxylin stained cell nuclei and the neuropil, and the pale background. Regions of the combined hematoxylin stained cell nuclei and the neuropil are to be separated based on the differences in their regional shapes. The segmentation results of real immunohistochemical images of brain microglia are provided and discussed. © 2008 Society for Imaging Science and Technology.*
[DOI: 10.2352/J.ImagingSci.Technol.(2008)52:4(040502)]

INTRODUCTION

Image segmentation is essential in the quantitative analysis of cytological images.[1–5] Nucleus segmentation that separates the nucleus regions from other part of the images can provide diagnostically important information such as the nucleus sizes and shapes.[6,7] Nucleus segmentation also enables the subsequent image analysis to be performed solely in the nucleus regions without the interference of the insignificant image background.[8,9] The quantitative image analysis of brain immunohistochemistry staining may detect early axonal damage in brain injuries.[10] Unlike the natural images that vary a great deal in color and contents, the cytological images acquired via microscopes on the specimens with immunohistochemistry staining[11–13] have relatively homogeneous appearances. There are few distinctive colors in an immunohistochemistry staining image. Because of the unevenness in staining process, there may be slight variations in colors and intensities among the pixels of the same organs or tissues. It is desired that the segmentation algorithm should be able to learn from the contents in images before classifying the groups of pixels or regions. In segmentation of lung cell images, Wu and Gil[14] presented an adaptive algorithm using a circular centroid combined with a linear centroid to approximate the vector clusters. In the immunohistochemical brain images, the vector clusters resemble more closely two linear clusters. If two linear centroids are introduced and trained on the source vectors, the two long clusters may be separated based on the trained centroids. The microglial cell pixels or vectors that form an individual cluster may be classified after clustering. Although the hematoxylin stained cell nuclei and the neuropil regions are not separable by vector clustering because they share the same vector space, they appear differently in their regional shapes. To make use of the difference between the regional shapes, we develop a region grow–shrink procedure to segment the hematoxylin stained cell nuclei.

Received Oct. 9, 2007; accepted for publication Apr. 7, 2008; published online Jul. 22, 2008.
1062-3701/2008/52(4)/040502/11/$20.00.

LINEAR CENTROIDS

Let a color digital image, $\tilde{\mathbf{X}}$, be defined in a rectangular image domain, $S=\{(n_1,n_2)\,|\,n_1=0,1,2,\ldots,N_1-1;\ n_2=0,1,2,\ldots,N_2-1\}$. The dimension of the image is $N_1\times N_2$. A pixel of the image at the location (n_1,n_2) that is composed of three RGB color components can be represented by a three-dimensional (3D) vector $\tilde{\mathbf{x}}(n_1,n_2)=(\tilde{x}_1(n_1,n_2),\tilde{x}_2(n_1,n_2),\tilde{x}_3(n_1,n_2))^t$, where $[\bullet]^t$ denotes the vector transpose. While it is helpful to regard the image array as an $N_1\times N_2$ matrix of 3D vectors such as $\tilde{\mathbf{X}}=[\tilde{\mathbf{x}}(n_1,n_2)]_{N_1\times N_2}$, it is convenient, for the purpose of analysis, to convert the matrix into a $1\times N$ row matrix of 3D vectors, $\tilde{\mathbf{V}}=[\tilde{\mathbf{v}}_n]_{1\times N}=[\tilde{\mathbf{v}}_{\varphi(n_1,n_2)}]_{1\times N}$, for $n=0,1,2,\ldots,N-1$ and $N=N_1N_2$, by a scanning operation $\varphi(\bullet,\bullet)$ that strings the elements row after row as $n=\varphi(n_1,n_2)=N_2n_1+n_2$, for $0\leqslant n_1<N_1$, $0\leqslant n_2<N_2$. Since each element in the $1\times N$ row matrix $\tilde{\mathbf{V}}$ is a 3D vector, $\tilde{\mathbf{V}}$ can also be considered as a $3\times N$ matrix with each column representing a vector corresponding to one pixel in the original image according to the mapping of φ.

Let the auto-covariance function be estimated by $\mathbf{C}=[c(i,j)]_{3\times 3}=(1/N)(\tilde{\mathbf{V}}-\tilde{\mathbf{u}}_{\tilde{\mathbf{V}}}\mathbf{u})(\tilde{\mathbf{V}}-\tilde{\mathbf{u}}_{\tilde{\mathbf{V}}}\mathbf{u})^t$, where the mean vector $\tilde{\mathbf{u}}_{\tilde{\mathbf{V}}}=(1/N)\Sigma_{n=0}^{N-1}\tilde{\mathbf{v}}_n$ and $\mathbf{u}=[1\ 1\ 1\ldots 1]_{1\times N}$. The auto-covariance matrix $\mathbf{C}$ is symmetric and nonnegative. We have

$\mathbf{C}\tilde{\mathbf{A}} = diag(\lambda_1, \lambda_2, \lambda_3)\tilde{\mathbf{A}}$, where $\tilde{\mathbf{A}} = [\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3]$ and the eigenvalues in decreasing order $\lambda_1 \geqslant \lambda_2 \geqslant \lambda_3$. Since the third eigenvalue is generally much smaller than the other two eigenvalues, we discard it to simplify the computations. Let $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2]$ and the matrix $\mathbf{V}$ is a transform by $\mathbf{V} = \mathbf{A}^t\tilde{\mathbf{V}}$. The matrix $\mathbf{V}$, size of $2 \times N$, can be considered as a row vector of two-dimensional (2D) vectors such as $\mathbf{V} = [\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_{N-1}]$. Thus, we have N two-dimensional source vectors, $\mathbf{v}_n = [v_n(1), v_n(2)]^t$, for $n = 0, 1, \ldots, N-1$. With a linear mapping to move the ranges of the vector components into the image range grid, we have the transformed images corresponding to the first two eigenvalues as $x_k(n_1, n_2) = v_{\varphi^{-1}(n)} \times (k)$, for $k = 1, 2$, and $n = 0, 1, 2, \ldots, N-1$. Figure 2 shows the two images after the linear transform by the two eigenvectors, $\mathbf{a}_1$ and $\mathbf{a}_2$, corresponding the two largest eigenvalues λ_1 and λ_2, respectively. Image (a) has much higher contrast since its corresponding eigenvalue is much larger.

In the images from immunohistochemical stains of brain microglia utilizing diaminobenzidene (DAB) chromogen and hematoxylin counterstain, the image contents display roughly three visually different colors. These slides display cell nuclei and processes that have significantly different colors. However, since all slides stained with this standard procedure are similar in color, their vectors are located in vicinities. For a large digital image with N pixels, the vectors in $\mathbf{V}$ may form large clouds if they are displayed in an image with intensities corresponding to the number of vectors in $v(1) - v(2)$ space as shown in Fig. 2(a) where $v(1)$ is the vertical axis and $v(2)$ is the horizontal axis. Unlike many cases where the clusters appear round, Fig. 2(a) shows clearly two separate clusters that appear stretched out to form long and narrow shapes. Since the clusters appear straight, we use two straight lines instead of two points as the centroids to approximate the clusters. Setting the goal function as the mean-squared error (MSE) of the approximation of the clusters with the lines, we minimize this goal by iteratively reclassifying and updating the vectors. Starting with the initial lines as in Figure 3(b), the final lines of the centroids are obtained as shown in Fig. 3(g) when the procedure converges. The vast majority of vectors are in the larger cluster called major cluster. We call the other smaller cluster the minor cluster. The minor cluster contains the vectors of DAB chromogen that appear brown in color in the original image.

To separate the two clusters, we need to define a boundary or a line so that vectors located on one side are classified into one cluster and vectors on the other side are classified into the other cluster. We rotate the image by the angle of the centroid line of the major cluster so that the line is parallel to one axis of the domain. The image after rotation is depicted in (i) that shows the major cluster is stretched along the vertical axis. The partitioning of the domain is shown in Fig. 3(i) where the vertical line separates the minor cluster first, and the second horizontal line partitions the remaining area of the major cluster into two subclusters. The algorithm will be described in details in the next section.

CLUSTERING WITH LINEAR CENTROIDS AND THE IMPLEMENTATION

The classification of pixels is carried out by classifying the two-dimensional vectors $\mathbf{v}_n$, for $n = 0, 1, \ldots, N-1$, related by the mapping $n = \varphi(n_1, n_2)$, where (n_1, n_2) is the coordinate of pixel in the image domain and n is the index of the 2D vector $\mathbf{v}_n$. The range of the components in vectors $\mathbf{v}_n$ may be out of the image intensity range of $[0, L-1]$, where the intensity level $L = 256$, because of the linear transform by the eigenvectors. For convenience in display as an image, storage as a matrix and simplicity in computation of the histogram, we use a dynamic mapping that shift and scale the ranges of both the components in the vectors to inside the image intensity range. Let $\xi^+(k) = \max\{v_n(k)\}$ and $\xi^-(k) = \min\{v_n(k)\}$, for $k = 1, 2$, where $\max\{\cdot\}$ and $\min\{\cdot\}$ select, respectively, the maximum and minimum numbers from the sets. The ranges of the two components are $[\xi^-(1), \xi^+(1)]$ and $[\xi^-(2), \xi^+(2)]$. Let $\xi = \max\{(\xi^+(1) - \xi^-(1)), (\xi^+(2) - \xi^-(2))\}$. To move the range of each component to the image grid, we apply to the components the linear mappings of $round(((L-1)/\xi)(v(k) - \xi^-(k)))$, $k = 1, 2$. For the convenience of analysis and efficient usage of storage, we still use the $\mathbf{v}_n = [v_n(1), v_n(2)]^t$ to represent the vectors after the mapping, such as $v_n(k) = round(((L-1)/\xi)(v_n(k) - \xi^-(k)))$, for $k = 1, 2$, and $n = 0, 1, 2, \ldots, N-1$. The corresponding images from de-stacking $\mathbf{v}_n$ by $x_k(n_1, n_2) = v_{\varphi^{-1}(n)}(k)$, for $k = 1, 2$, and $n = 0, 1, 2, \ldots, N-1$, are shown in Fig. 2, where the image (a) shows higher contrast in intensities since its corresponding eigenvalue is larger.

Initializing zeros to $\mathbf{h}$, an image of size $L \times L$. Each element in $\mathbf{h}$ is a 2D vector or a paired variable. The 2D histogram of the vectors is computed by one loop as the following:

$$\text{for } n = 0 \text{ to } N-1 \text{ do } \mathbf{h}(v_n(1), v_n(2)) = \mathbf{h}(v_n(1), v_n(2)) + 1.$$

Since the range of components in $\mathbf{v}(n)$ is limited in $[0, L-1]$, the squared domain of $\mathbf{h}$ is side of $L-1$. Fig. 3(a) show the 2D histogram of the images in Fig. 2, where the vertical axis corresponds to the image in Fig. 2(a) and the horizontal axis corresponds to the image in Fig. 2(b). Higher intensity means higher number of identical vectors in the set $\{\mathbf{v}_n\}$.

Our first objective is to separate the two clusters that appear long and narrow in shapes. Using two straight lines as their respective centroids, we can significantly simplify the clusters with only a few coefficients. Let $I = 2$ be the number of clusters. Each line can be defined by a pair of coefficients α and ρ. Let the ith line be defined by $v(1)\cos(\alpha_i) + v(2)\sin(\alpha_i) - \rho_i = 0$, for $1 \leqslant i \leqslant I$. The distance between the point $\mathbf{v}_n$ and the ith line centroid will be $d_{i,n} = |v_n(1)\cos(\alpha_i) + v_n(2)\sin(\alpha_i) - \rho_i|$. Let S_i be a set composed of all the closest source vectors, i.e., $S_i = \{\mathbf{v}_n | \min^{-1}\{d_{j,n}\} = i\}$, where the inverse minimum operation $\min^{-1}_{1 \leqslant j \leqslant I} d_{j,n}$ yields the index j, instead of the distance, corresponding to the smallest distance among the I distances

$d_{j,n}$, for $1 \leqslant j \leqslant I$. If $\min^{-1}_{1 \leqslant j \leqslant I} d_{j,n} = i$, we assign the vector $\mathbf{v}_n$ into the cluster represented by the ith line. The set S_i consists of the source vectors that are closer to the ith line centroid than to any of the other $I-1$ lines. The cost of classifying vectors in S_i is measured by the mean-squared error as $\varepsilon_i^2 = (1/M_i)(\Sigma_{n:\mathbf{v}_n \in S_i}(v_n(1)\cos(\alpha_i) + v_n(2)\sin(\alpha_i) - \rho_i)^2)$. The average cost of classification of all source vectors is measured by $\varepsilon^2 = (1/N)\Sigma_{i=1}^{I}(M_i\varepsilon_i^2)$. To update the line centroids for the source vectors in S_i, we recomputed the line parameters so that the cost of ε_i^2 is minimized. The parameters of the straight line, ρ_i and α_i, for $1 \leqslant i \leqslant I$, can be obtained based on the source vectors according to the Eqs. (A1) and (A2) in the Appendix. The difference between the two possible solutions for α_i is a constant angle of $\pi/2$. The selection of α_i between the two values should be determined by the cluster of the source vectors since only one is the right choice. The correct α_i renders the line fitting more properly to the cluster and results in a significantly lower classification cost, ε_i^2.

We have shown that for a given set S_i we can find the best fitting line, determined by the pair (α_i, ρ_i), that minimizes the approximation error ε_i^2. The average error for all vectors $\varepsilon^2 = (1/N)\Sigma_{i=1}^{I}(M_i\varepsilon_i^2)$ will be smaller if we reclassify the source vectors so that S_i contains the vectors that are located closer to the line of (α_i, ρ_i) than any of the other lines. With such a modified partitioning $\{S_i | 1 \leqslant i \leqslant I\}$ that reduces the error ε^2, we can still find an even better fitting line in each set of S_i according to the Eqs. (A1) and (A2). Thus, we can develop an iterative procedure to repetitively update the lines and reassign the vectors. Since the error decreases with an additional iteration, the feedback procedure converges. The iterative algorithm is summarized as the following:

Step 0. *Initialization:*

Given the vectors $\mathbf{v}_n$, for $n = 0, 1, \ldots, N-1$; I, the number of centroid lines; initial line parameters $\alpha_i^{(0)}$ and $\rho_i^{(0)}$ for $1 \leqslant i \leqslant I$; the initial iteration index $m = 0$, a large number $(\varepsilon^2)^{(-1)} = 1.0e30$ and a small number $\delta = 0.001$.

Step (1): *Nearest neighbor reassignment:*

Test the distance of each vector to each of the centroid lines. If the distance to the line of $(\alpha_i^{(m)}, \rho_i^{(m)})$, assign or reassign the vector to S_i. The sets are recalculated by

$$S_i^{(m)} = \left\{ \mathbf{v}_n \middle| \min^{-1}_{1 \leqslant j \leqslant 1} \{ |v_n(1)\cos(\alpha_j^{(m)}) + v_n(2)\sin(\alpha_j^{(m)}) - \rho_j^{(m)}| \} = i \right\}.$$

Step (2): *Termination criterion:* Calculate the average error $(\varepsilon^2)^{(m)} = \frac{1}{N}\Sigma_{i=1}^{I}\Sigma_{n:\mathbf{v}_n \in S_i^{(m)}}(v_n(1)\cos(\alpha_i^{(m)}) + v_n(2) \times \sin(\alpha_i^{(m)}) - \rho_i^{(m)})^2$. If $((\varepsilon^2)^{(m-1)} - (\varepsilon^2)^{(m)})/(\varepsilon^2)^{(m)} < \delta$ then terminate the iterations and output the classification $S_i^{(m)}$ and the parameters $\alpha_i^{(m)}$ and $\rho_i^{(m)}$, otherwise $m = m+1$ and go to the next step.

Step (3): *Update of the centroid lines:*

Updating the centroids by adjusting the line parameters $\alpha_i^{(m)}$ and $\rho_i^{(m)}$. Try $\alpha_i^{(m)} = \frac{1}{2}\tan^{-1}(2(a_1b_1 - c)/((a_1^2 - a_2) - (b_1^2 - b_2)))$, where $a_1 = (1/\mathbf{M}_i^{(m-1)})\Sigma_{n:\mathbf{v}_n \in S_i^{(m-1)}} v_n(1)$, $b_1 = (1/\mathbf{M}_i^{(m-1)})\Sigma_{n:\mathbf{v}_n \in S_i^{(m-1)}} v_n(2)$, $a_2 = (1/\mathbf{M}_i^{(m-1)})\Sigma_{n:\mathbf{v}_n \in S_i^{(m-1)}} v_n^2(1)$, $b_2 = (1/\mathbf{M}_i^{(m-1)})\Sigma_{n:\mathbf{v}_n \in S_i^{(m-1)}} v_n^2(2)$, and $c = (1/\mathbf{M}_i^{(m-1)})\Sigma_{n:\mathbf{v}_n \in S_i^{(m-1)}} v_n(1)v_n(2)$. If $|\alpha_i^{(m)} - \alpha_i^{(m-1)}| > \pi/4$, recompute $\alpha_i^{(m)} = \frac{1}{2}\tan^{-1}(2(a_1b_1 - c)/((a_1^2 - a_2) - (b_1^2 - b_2))) + \pi/2$. Let $\rho_i^{(m)} = a_1\cos(\alpha_i^{(m)}) + b_1\sin(\alpha_i^{(m)})$. If $\rho_i^{(m)} < 0$, then adjust $\rho_i^{(m)} = -\rho_i^{(m)}$ and $\alpha_i^{(m)} = \alpha_i^{(m)} - \pi$. Return to step 1 for reassignment of the source vectors based on the new line centroids.

Separation of the Two Clusters

In a brain immunohistochemical image, the microglial cell cytoplasm that appears in brown color with DAB staining has a very small area. The microglial cell cytoplasm occupies less than 1% of the total image area. The major cluster dominates and minor cluster can be considered as formed by some stray vectors. Thus, the boundary separating the two clusters should be determined based on the major cluster. The boundary is a line parallel to the major centroid line. We rotate the clusters so that the major centroid line is parallel to one of the two axes. Since the two eigenvalues are usually significantly different and the minor cluster is very small, the centroid line of the major cluster should not be far from parallel to one of the axes.

With a rotation of axes by the angle of α_1, the centroid line of the major cluster will appear parallel to the vertical axis in the new coordinate system. Let the center of image be denoted by (T_1, T_2), where $T_1 = round(N_1/2)$ and $T_2 = round(N_2/2)$. The new vectors, $\mathbf{v}_n' = (v_n'(1), v_n'(2))^t$, for $n = 0, 1, 2, \ldots, N-1$, after the rotation about the image center point of (T_1, T_2) is derived by $v_n'(1) = (v_n(1) - T_1)\cos(\alpha_1) - (v_n(2) - T_2)\sin(\alpha_1) + T_1$ and $v_n'(2) = (v_n(1) - T_1)\sin(\alpha_1) - (v_n(2) - T_2)\cos(\alpha_1) + T_2$. Fig. 3(a) shows the clusters of $\mathbf{v}_n$, for $n = 0, 1, 2, \ldots, N-1$, while Fig. 3(h) shows the clusters of $\mathbf{v}_n'$, for $n = 0, 1, 2, \ldots, N-1$, after rotating the clusters in (a) around the image center so that the major cluster becomes vertical to the horizontal axis.

The histogram of the new vectors can be calculated by such a loop as: for $n = 0$ to $N-1$ do

$$\mathbf{h}'(v_n(1), v_n(2)) = \mathbf{h}'(v_n(1), v_n(2)) + 1.$$

The horizontal distribution of the vectors $\mathbf{v}_n'$, for $n = 0, 1, 2, \ldots, N-1$, is computed by $\mathbf{h}_1(l) = (1/N)\Sigma_{j=0}^{L-1}\mathbf{h}(j, l)$, for $l = 0, 1, 2, \ldots, L-1$. Figure 5(a) displays the reduced histogram $\mathbf{h}_1(l)$ from the 2D histogram in Figure 3(h). It is observed that the waveform is smooth and nearly symmetric except a long and thin tail on the left side for the minor cluster. If we can assume the symmetric property for the major cluster distribution, then we can find the cutoff point at the left side of the waveform based on the right cutoff point which is easier to find because there is no interference from unwanted vectors in this area. Let $h_{1,\max} = \{\mathbf{h}_1(l) | l = 0, 1, 2, \ldots, L-1\}$. Find the two points, l_1 and l_2, for which the function $\mathbf{h}_1(l)$ crosses over the half line of its peak. Find l_1 such that $\mathbf{h}_1(l_1) \leqslant h_{1,\max}/2$ and $\mathbf{h}_1(l_1 + 1) > h_{1,\max}/2$, and l_2

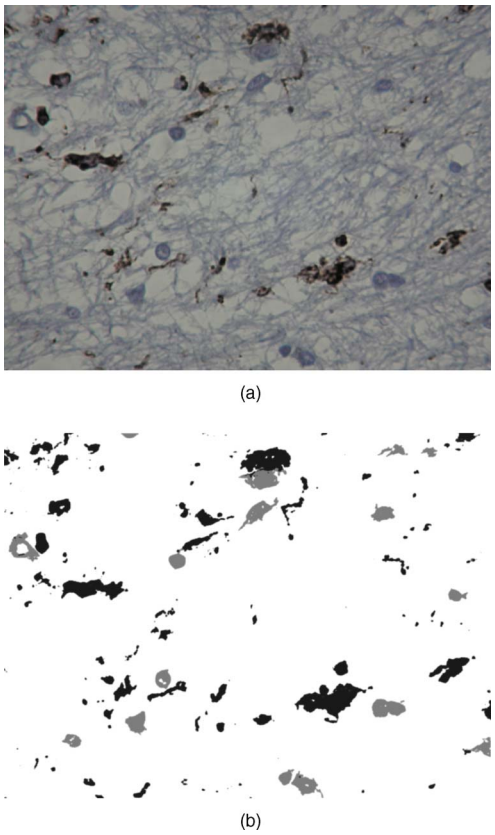


Figure 1. (a) An original discrete color image of brain microglia stained with an antibody to CD68 antigen, with diaminobenzidene as a chromogen and hematoxylin counterstain; (b) the segmented image, black for stained activated microglia, gray for cell nuclei, and white for cell neuropil.

such that $\mathbf{h}_1(l_2-1)>h_{1,\max}/2$ and $\mathbf{h}_1(l_2)\leqslant h_{1,\max}/2$. The center point is then $(l_1+l_2)/2$. The accumulative distribution is $\hat{\mathbf{h}}_1(l)=\Sigma_{j=0}^{l}\mathbf{h}_1(j)$, for $0\leqslant l<L$. $\hat{\mathbf{h}}_1(l)$ is an ascending function and limited between 0 and 1. The upper limit for the major cluster is where $\hat{\mathbf{h}}_1(l)$ approaches one. If $\hat{\mathbf{h}}_1(\tau_1^+)\leqslant 0.999$ and $\hat{\mathbf{h}}_1(\tau_1^+ +1)>0.999$, the intensity τ_1^+ is considered to be the upper limit for the major cluster as seen in Fig. 5(a). The boundary is at τ_1^- which is the mirror point of τ_1^+ about the center at $(l_1+l_2)/2$, such as $\tau_1^-=(l_1+l_2)/2-(\tau_1^--(l_1+l_2)/2)$ or $\tau_1^-=l_1+l_2-\tau_1^+$. The corresponding boundary line in 2D clusters is shown in Figure 3(i) as the vertical white line that separates the two clusters.

The vectors classified in the minor cluster correspond to brown-colored microglial cell cytoplasm pixels. The major cluster contains the remaining vectors including the cells that are stained in blue color. To separate the blue staining cells whose vectors occupy the lower area in the major cluster, we use a straight line to partition the major cluster as the horizontal line shown in Figure 3(i). The procedure is similar to the procedure above finding the vertical boundary line. The vertical distribution of the vectors in the major cluster is computed by $\mathbf{h}_2(l)=(1/N)\Sigma_{j=\tau}^{L-1}\mathbf{h}(l,j)$, for $l=0,1,2,\ldots,L-1$. Note that in the sum j starts at τ_1^- instead of zero. Figure 5(b) displays the reduced histogram $\mathbf{h}_2(l)$. Let $h_{2,\max}=\{\mathbf{h}_2(l)\,|\,l=0,1,2,\ldots,L-1\}$. Find the two points, l_1 and l_2, for which the function $\mathbf{h}_2(l)$ crosses over the half line of its peak. Find l_1 such that $\mathbf{h}_1(l_1)\leqslant h_{1,\max}/2$ and $\mathbf{h}_1(l_1+1)>h_{1,\max}/2$, and l_2 such that $\mathbf{h}_1(l_2-1)>h_{1,\max}/2$ and $\mathbf{h}_1(l_2)\leqslant h_{1,\max}/2$. If there are multiple values, choose the smallest for l_1 and the largest for l_2. The center point is again $(l_1+l_2)/2$. The accumulative distribution is $\hat{\mathbf{h}}_2(l)=\Sigma_{j=0}^{l}\mathbf{h}_2(j)$, for $0\leqslant l<L$. The upper limit for the major cluster is where the ascending function $\hat{\mathbf{h}}_2(l)$ approaches one. If $\hat{\mathbf{h}}_2(\tau_2^+)\leqslant 0.999$ and $\hat{\mathbf{h}}_2(\tau_2^+ +1)>0.999$, the intensity τ_2^+ is considered to be the upper limit for the major cluster as seen in Figure 5(a). The boundary is at τ_2^- which is the mirror point of τ_2^+ about the center at $(l_1+l_2)/2$, such as $\tau_2^-=(l_1+l_2)/2-(\tau_2^--(l_1+l_2)/2)$ or $\tau_2^-=l_1+l_2-\tau_2^+$. The corresponding boundary line in 2D clusters is shown in Fig. 3(i) as the horizontal white line that separates the two clusters.

The two straight lines, one vertical and the other horizontal, partition the whole space into three rectangular regions. Vectors are classified into one of the three groups according to the regions where they are located. We have the segmentation

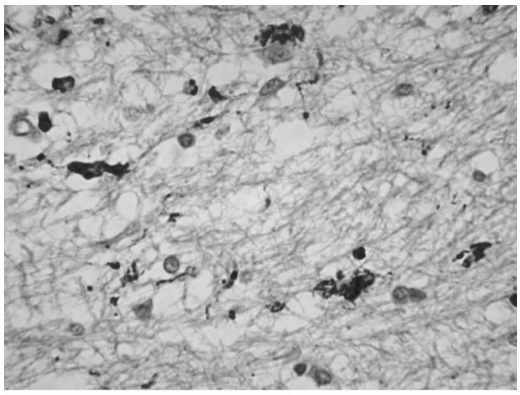
$$\Xi(n_1,n_2)=\begin{cases}0, & \text{for } \nu'_{\varphi(n_1,n_2)}(2)<\tau_1^-;\\ 1, & \text{for } \nu'_{\varphi(n_1,n_2)}(2)\geqslant\tau_1^- \text{ and } \nu'_{\varphi(n_1,n_2)}(1)<\tau_2^-;\\ 2, & \text{otherwise.}\end{cases}$$

The image segmentation, $\Xi(n_1,n_2)$ for $(n_1,n_2)\in S$, corresponding to the partition in Figure 3(i) is displayed by the three-level image in Figure 6 where black intensity stands for the vectors in the rectangular region on the left in Fig. 3(i), gray intensity for the lower right rectangular region, and white intensity for the upper right rectangular region.

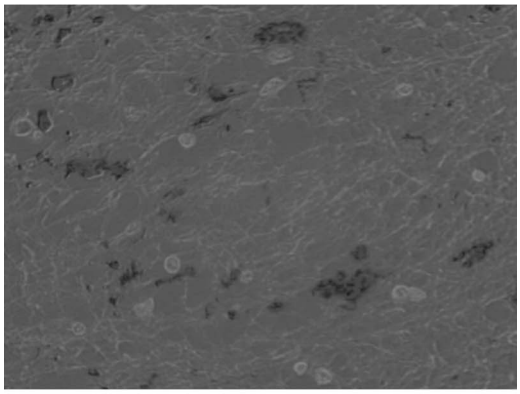
The dark intensity in Fig. 6 represents the segmented DAB stained microglal cytoplasm pixels. Since the minor cluster is clearly separable from the major cluster, the microglial cell pixels are classified accurately as observed when comparing the dark regions with the brown regions in the original image in Figure 1(a). However, the gray intensity regions for $\Xi(n_1,n_2)=1$ in Fig. 6 consist of hematoxylin stained cell nuclei and processes appear in relatively bluish color. They are not separable simply based on their colors. To discriminate the blue appearing nuclei from blue appearing processes, we develop an approach based on their differences in shapes in the next section.

SEPARATION OF HEMATOXYLIN STAINED CELLS FROM NEUROPIL REGIONS

The cells appear blue in the original image, so are the neuropil (cell processes) regions. The similarity in color between the cells and the neuropil (cell processes) make the discrimination between them difficult. As observed in segmentation in Fig. 6, the gray areas consist of not only the cells but also a portion of the neuropil regions. It is also observed that there are differences in geometric shapes between the cell regions and neuropil regions. The cells appear almost round in shape although they there are occasionally holes inside in the segmented gray leveled regions, while the segmented neuropil regions appear straw-like, slim, and long in variant



(a)



(b)

Figure 2. Transformed image. (a) Image corresponding to the largest eigenvalue; (b) image corresponding to the second largest eigenvalue.

shapes and sizes. If we repetitively shrink the gray regions with morphologic erosions, we can expect that the slim neuropil regions vanish first.

The erosion process is a cascade of the same erosion operators. Let the input to an erosion operator be $g_i(n_1,n_2)$ and the output $g_o(n_1,n_2)$. The input to the first erosion operator is the image corresponding to the gray intensity defined as $g'(n_1,n_2)=\begin{cases}1, \text{ if } \Xi(n_1,n_2)=1;\\ 0, \text{ otherwise.}\end{cases}$ and the output from the last erosion operator is denoted by $g''(n_1,n_2)$. Let the erosion kernel be a circular region of small radius r. With the input $g_i(n_1,n_2)$, the erosion of the image[15] is obtained by the following procedure:

(1) Initialize $g_o(n_1,n_2)=g_i(n_1,n_2)$, for $(n_1,n_2)\in S$;
(2) If $g_i(n_1,n_2)=0$, for $(n_1,n_2)\in S$, let $g_o(n_1+m_1,n_2+m_2)=0$, for $m_1^2+m_2^2<r^2$.

The above erosion shrinks the foreground areas at their edges. Since an irregular region that has higher perimeter to area ratio shrinks more than a circular region on the area percentage basis, the irregular and thin neuropil regions may disappear earlier than the circular cell regions if a repetitive erosion is applied. Thus, if we stop the shrinking process before the cell regions disappear based on the average cell size which is usually known roughly, we can reduce the neuropil regions significantly.

A reverse process to the repetitive erosions is to restore the cell regions. Corresponding to each erosion in the shrinking procedure is a constrained dilation. Since many neuropil regions may disappear completely after the cascade of erosions, they are not to be recovered in the reverse process. Let the input to a constrained dilation operator be $g_i(n_1,n_2)$ and the output $g_o(n_1,n_2)$. The input to the first dilation operator is $g''(n_1,n_2)$, the output of the final erosion operator. The constrained dilation is as follows.

(1) Initialize $g_o(n_1,n_2)=g_i(n_1,n_2)$, for $(n_1,n_2)\in S$;
(2) If $g_i(n_1,n_2)=1$, for $(n_1,n_2)\in S$, let $g_o(n_1+m_1,n_2+m_2)=1$, for $m_1^2+m_2^2<r^2$;
(3) Let $g_o(n_1,n_2)=g_o(n_1,n_2)^*g'(n_1,n_2)$, for $(n_1,n_2)\in S$.

RESULTS

Figure 1(a) is an original image of 24 bit RGB color acquired from a section of brain stained immunohistochemically for microglial cell cytoplasm using an antibody to CD68 antigen, with diaminobenzidene as a chromogen and color with hematoxylin staining. The size of the digital image is $N_1\times N_2=1600\times 1200$. The brain immunohistochemical image contains microglial cell cytoplasm that appears in brown color, hematoxylin stained cell nuclei in blue color, and neuropil regions that appear in either light blue or gray.

The first step is to compute the eigenvalues and eigenvectors of the 3×3 auto-covariance matrix from the stacked row of vectors. Multiplying the eigenvector matrix, we have the new vectors, which have total energy concentrated in fewer components according to the eigenvalues. In this particular case, the three eigenvalues are 1537.5, 111.2, and 21.5, respectively. The first two images corresponding to the two largest eigenvalues are displayed in Figures 2(a) and 2(b), respectively. The third eigenvalue is very small comparing to the first two and thus its corresponding component is discarded to simplify the computational complexity without large loss. The two-dimensional vector distribution is displayed in Fig. 3(a) where the vertical axis is for the component corresponding to the larger eigenvalue and the horizontal axis is for the other component. Since the vertical component is much larger, we expect the clusters to spread approximately along the vertical axis. The perpendicular lines to the cluster lines should be approximately zero. Thus, we let the initial angles of the two centroid lines be zero, i.e., $\alpha_1^{(0)}=0$ and $\alpha_2^{(0)}=0$. The initial lines are parallel to the vertical axis as shown in Figure 3(b). To set the two initial lines on separate side of the vector clusters on the horizontal axis so that the vectors located on both ends are classified correctly in different groups, the two distances of the lines from the origin are initialized as $\rho_1^{(0)}=170$ corresponding to the line on the right of the clusters and $\rho_2^{(0)}=20$ corresponding to the other line on the right side of the clusters as shown in Figure 3(b). Starting with initial centroids, the iterative procedure repetitively reassigns the source vectors and updates centroid lines until the termination test is satisfied. Figs. 3(c)–3(f) show the updated centroid lines after 1, 2, 3, and 4 iterations, respectively. As we can see from the image se-

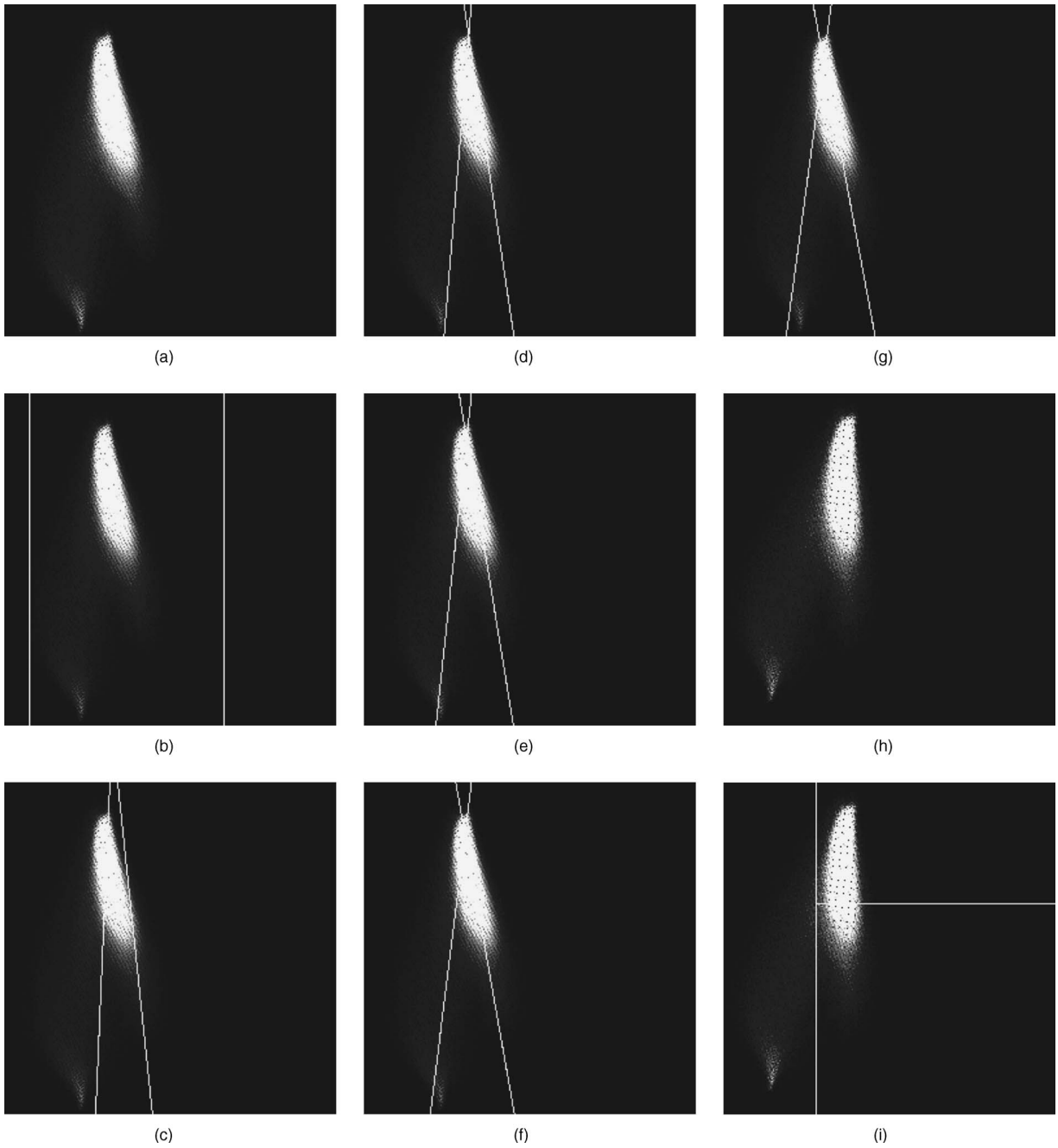


Figure 3. (a) Two-dimensional vector distribution; (b) initial centroid lines superimposed in the vector distribution in (a); (c) after one iteration; (d) after two iterations; (e) after three iterations; (f) after four iterations; (g) after 23 iterations when no changes happen with additional iterations; (h) rotate the image (a) by the angle of α_1; and (i) the derived boundaries partitioning the space.

quence, the centroid lines are adjusted gradually toward the centers of the clusters. Fig. 3(g) shows the final centroids after 23 iterations when no changes happen with any additional iteration. Figure 4 shows the traces of the line centroids with the increasing number of iterations. The changes or updates in both α and ρ are large in the first few iterations and then decrease gradually. The converged centroid lines are $\alpha_1=0.1927$, $\rho_1=115.9$, $\alpha_2=-0.1414$, and $\rho_2=45.83$. We rotate the coordinate plane of Fig. 3(g) by the angle of α_1 about the center of the image to have the major cluster parallel to the vertical axis as shown in Fig. 3(h) so that the partitioning boundary between the two clusters becomes a simple vertical line whose location is determined by the distributions of the source vectors. Accumulating the vectors vertically in Fig. 3(h), we have the one-dimensional vector distribution, $\mathbf{h}_1$, as shown Fig. 5(a). We can see that substantial number of vectors are located in a narrow area resulting in a large spike between l_1 and l_2 which are the horizontal locations of middle points between 0 and the top of the waveform. The area of the microglial cell cytoplasm that appears in brown color in the original image is usually less than 2% of the total area. The small number of brown

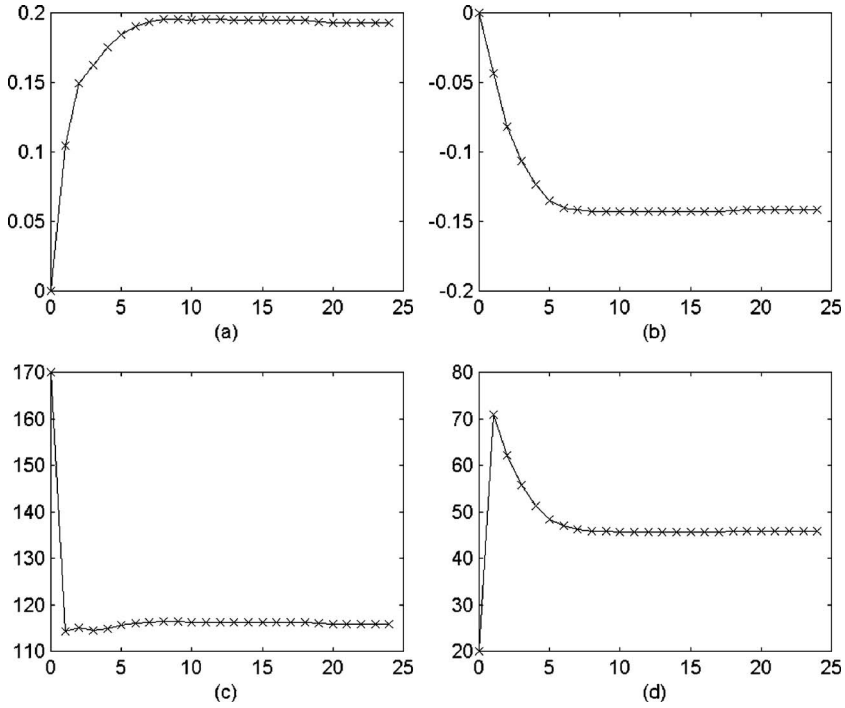


Figure 4. Convergence of the centroid lines vs m, the number of iterations. (a) $\alpha_1^{(m)}$ vs m, (b) $\alpha_2^{(m)}$ vs m, (c) $\rho_1^{(m)}$ vs m, and (d) $\rho_2^{(m)}$ vs m.

colored pixels corresponds to the long tail of low values of the distribution function of $\mathbf{h}_1$ at the left side of the large spike. If the tail is excluded, the remaining distribution function is approximately symmetric and smooth. This feature is general because the pixels in cells and the neuropil regions that usually occupy more than 99% of area in the image share the similar color. Assuming the symmetry property, we obtain the value $\tau_1^- = 72$. Thus, we have the partitioning boundary between the two clusters shown as the vertical line in Fig. 3(i). Similarly, the partitioning line regions that are perpendicular to the long axis of the approximately elliptical shaped cluster of the cell and neuropil regions are based on the vector distributions of $\mathbf{h}_2$ in Fig. 5(b) and shown as the horizontal line in Fig. 3(i). Thus, the two straight lines in Fig. 3(i) partition the domain into three separate regions.

Fig. 6 shows the segmentation according to the partitioning lines in Fig. 3(i). The three intensities in the segmentation image represent the classifications of vectors in the respective three regions in Fig. 3(i). In Fig. 6, black, gray and white pixels correspond to the rectangular regions on the left, in Fig. 3(i), gray to the lower right rectangular region, and white to the upper right rectangular region. From the image $\Xi(n_1, n_2)$ in Fig. 6, we can obtain the segmentation $g'(n_1, n_2)$. The radius of the circular structure elements should be a small value. In our experiments, the radius is selected to be 2.5. If a larger value is selected, the number of erosions and dilations should be smaller but each operation consumes much longer computational time. With each erosion, the nuclear regions are shrunk by 5 in diameters. Considering the size of the nuclei that are generally larger than 30 in pixels in diameters in the original images used in this article, a series of fewer than six successive erosion operations may be applied in order to retain portion of nuclei followed by a series of the same number of constrained dilations. Figure 1(b) shows the segmented image where black intensity stands for stained activated microglia, gray for the cell nuclei, and white for the cell neuropil. Figure 7 shows another experimental result where (a) is the original image and (b) is the resulting segmentation where black stands for stained activated microglia, gray for the cell nuclei, and white for the cell neuropil. In visual comparisons, the segmentations match their original images.

Recently, Ruifrok and Johnston[16] have presented a color deconvolution algorithm that transforms the image RGB components into a set of nonorthogonal components corresponding to the major colors in the images. In Figure 8, we show the best segmentation by the manual thresholding based on the color components from the deconvolution algorithm using their built-in vectors for hematoxylin and DAB separations. Each of the DAB and hematoxylin components after color deconvolution is a gray level image. To determine the microglia areas, a threshold is applied on the DAB color component. The threshold is adjusted manually so that the best segmentation, shown in Fig. 8(a), based on visual judgment, is achieved. It is observed that the segmented microglia regions in Fig. 8(a) and Fig. 1(b), both in black, match well. While the results for microglia regions from both algorithms are comparable, the proposed algorithm has several advantages. The proposed algorithm is automatic because it processes the images based on their statistical properties and correlations without the inputs of users. Unlike the color deconvolution method that needs a set of fixed and preset conversion vectors, the proposed algorithm is adaptive to the images and therefore tolerant to

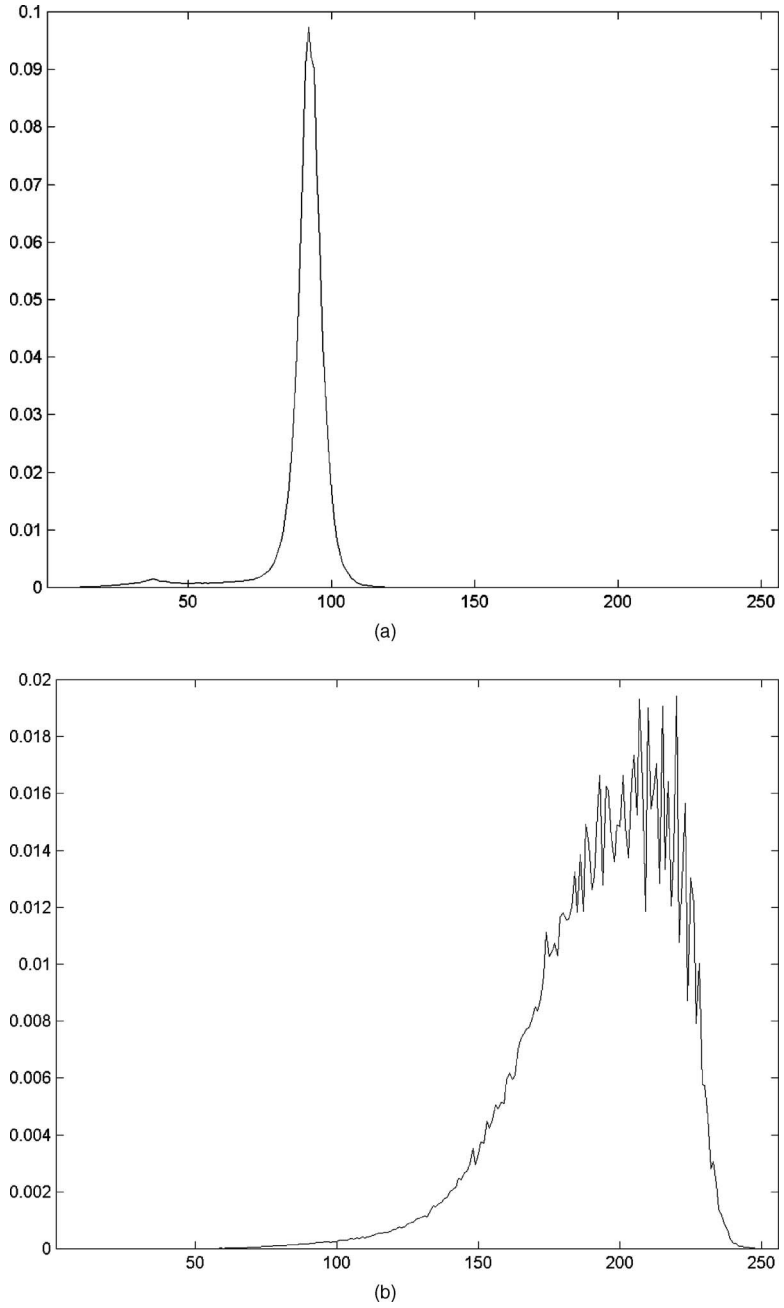


Figure 5. (a) $h_1(l)$, (b) $h_2(l)$.

the staining variations. Fig. 8(b) is the best segmentation result achieved by manually adjusting the threshold for the hematoxylin color component from the color deconvolution method. Since the hematoxylin stained cell nuclei and the neuropil have a similar color, they are not separable with a simple threshold on the hematoxylin color component from the color deconvolution method. Our proposed algorithm separates the nuclei based on their geometrical shapes as shown in gray in Fig. 1(b).

Wu and Gil[14] have proposed a clustering algorithm using mixed point and line centroids for segmentation of images of lung cells. The point centroid is appropriate for the blood cells in lung cell images since the vectors from blood cells are located in a small area. However, the vectors of the brown microglia pixels in a brain immunohistochemistry image usually occupy a large oblong area that is more appropriately approximated by a line instead of a point. For comparison, we apply the algorithm to the original image in Fig. 1(a) and show the results in Fig. 9. Fig. 9(a) shows the cluster image in Fig. 3(a) superimposed by the initial point centroid at the left-bottom corner as well as the initial vertical centroid line. Fig. 9(b) shows the centroids after eight iterations when the procedure converges. Figs. 9(c) and 9(d) show the respective classifications of the vectors. As we can see, there is a significant portion of vectors in the lower cluster of microglia in Fig. 9(c), which are misclassified to the other cluster in Fig. 9(d). The corresponding segmentation of microglia regions is shown in Fig. 9(e). Comparing it

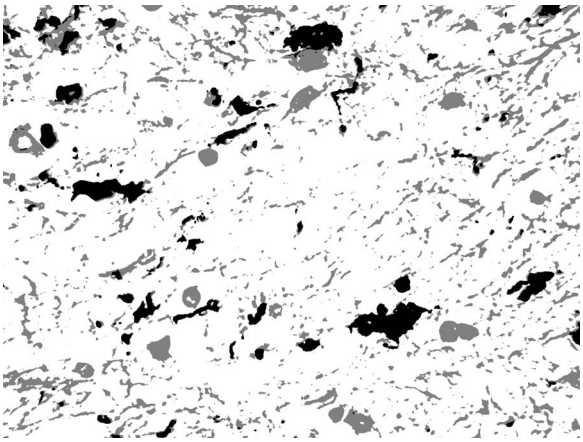


Figure 6. Segmentation by the partitioning lines in Figure 3(i). Three intensities represent the vectors in the respective three regions in Figure 3(i). Black corresponds to the rectangular region on the left in Figure 3(i), gray to the lower right rectangular region, and white to the upper right rectangular region.

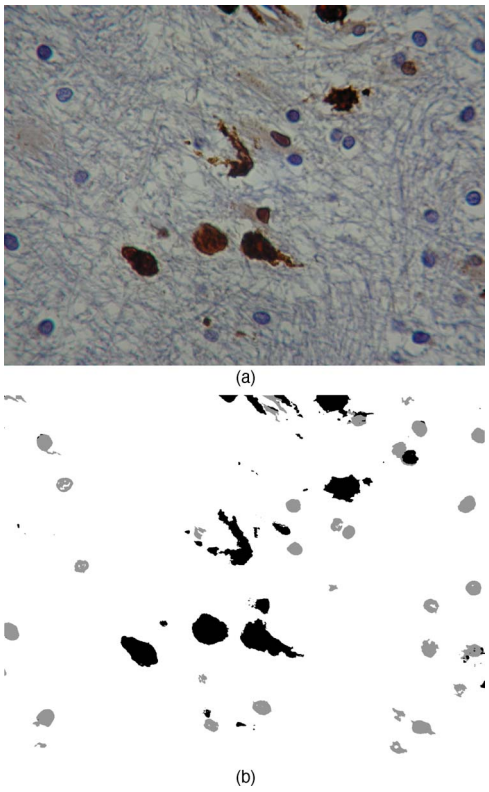


Figure 7. (a) Original image, (b) segmentation of the image in (a).

to the microglia regions in black in Fig. 1(b), we observe that some microglia pixels are lost.

CONCLUSIONS

We have described a segmentation algorithm for brain immunohistological images that have few different colors. Since the cell nuclei and neuropil that occupy an overwhelmingly large portion of image have similar bluish color in the counter-stained brain immunohistological images while activated microglia that occupies a tiny portion of the image

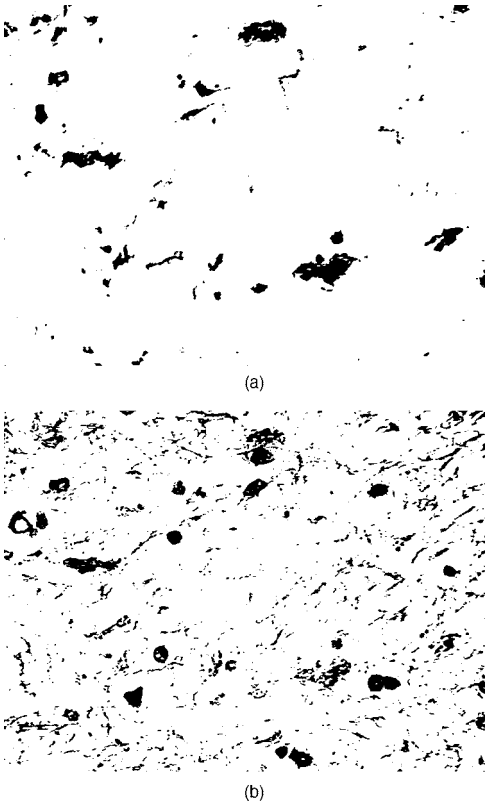


Figure 8. Segmentation based on the color deconvolution,(see Ref. 16) vectors chosen for Haematoxylin and DAB separation. (a) Manual thresholding of DAB component, (b) manual thresholding of haematoxylin component.

has a different color, the cluster of the cell nuclei and neuropil pixels is much denser and larger than the cluster of activated microglia pixels. The orthogonal transform based on the eigenvectors of the image pixel vectors is mostly determined by the larger cluster of the cell nuclei and neuropil. The prolonged cluster can always be placed along a particular direction based on the eigenvalues. An iterative learning procedure is developed to obtain the centroid lines, which are then used to partition the vector domain into regions of the activated microglia and the rest. The hematoxylin stained cell nuclei and the neuropil are segmented in the subsequent procedure based on the difference between their geometrical shapes. Figures show both the learning and segmentation processes for the real brain immunohistochemical images. The segmentation results are illustrated to show the effectiveness of the algorithm for visual comparisons. The algorithm may be applicable to images of other organs if they satisfy the assumptions that one cluster takes the overwhelmingly large portion (>95% for instance), but this remains for future studies.

ACKNOWLEDGMENTS

This work was supported in part by Grant No. R24MH59724 (The Manhattan HIV Brain Bank) from the National Institutes of Health.

APPENDIX

The MSE for the vectors in S_i approximated by the line centroid is

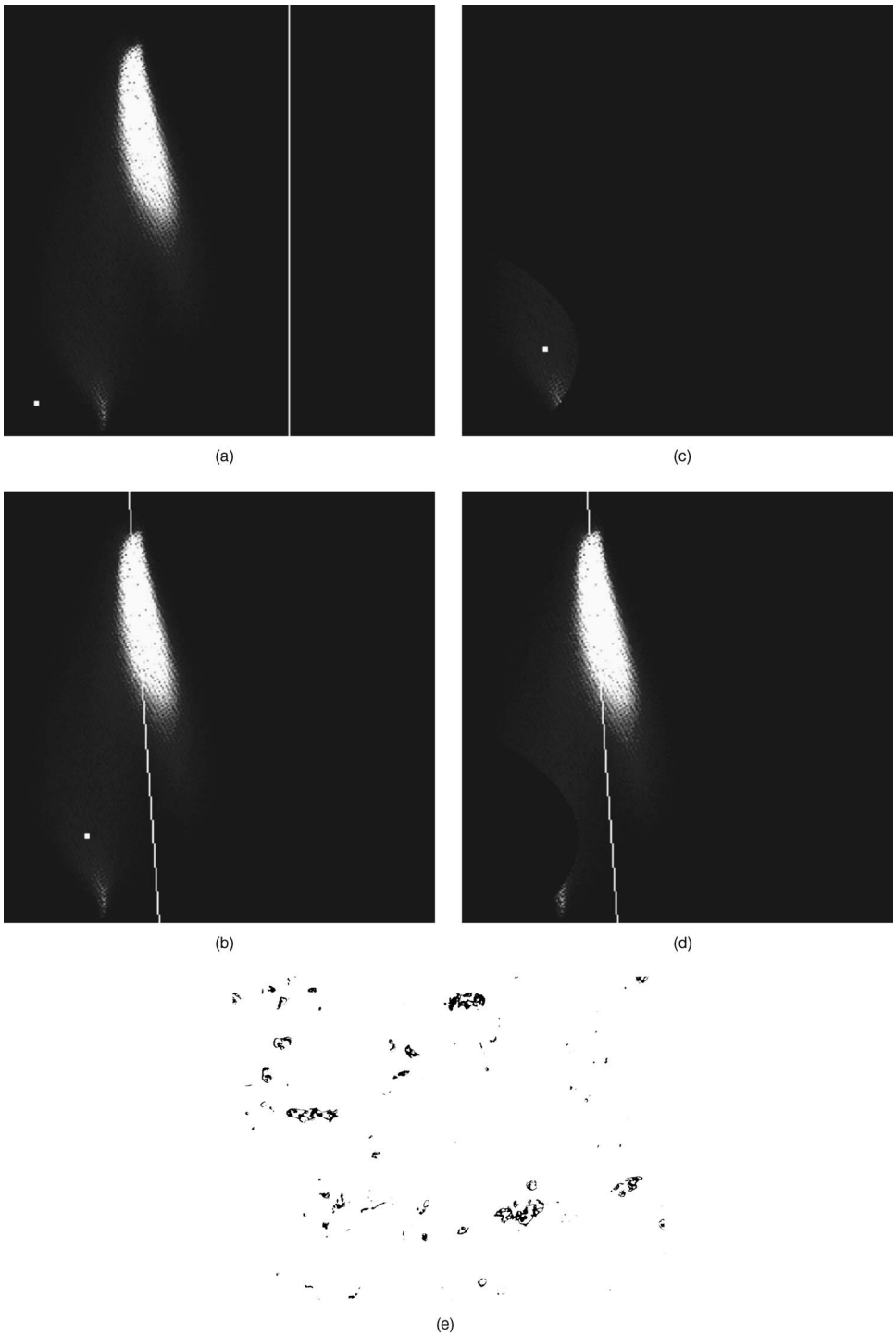


Figure 9. Results of the linear clustering with both point and linear centroids.(see Ref. 14) (a) The initial point centroid at the left-bottom corner and the vertical linear centroid in the middle, (b) centroids after eight iterations when procedure converges, (c) vectors near the point centroid classified to activated microglia, (d) the remaining vectors, and (e) segmentation of activated microglia corresponding to the vectors in (c).

$$\varepsilon_i^2 = \frac{1}{M_i}\left(\sum_{n:\mathbf{v}_n \in S_i} (v_n(1)\cos(\alpha_i) + v_n(2)\sin(\alpha_i) - \rho_i)^2\right).$$

We have

$$\frac{\partial \varepsilon_i^2}{\partial \rho_i} = \frac{-2}{M_i}\left(\sum_{n:\mathbf{v}_n \in S_i} (v_n(1)\cos(\alpha_i) + v_n(2)\sin(\alpha_i) - \rho_i)\right).$$

If

$$\frac{\partial \varepsilon_i^2}{\partial \rho_i} = 0, \quad \rho_i = \frac{1}{M_i}\left(\sum_{n:\mathbf{v}_n \in S_i} (v_n(1)\cos(\alpha_i) + v_n(2)\sin(\alpha_i))\right) \text{ or } \rho_i = \left(\frac{1}{M_i}\sum_{n:\mathbf{v}_n \in S_i} v_n(1)\right)\cos(\alpha_i) + \left(\frac{1}{M_i}\sum_{n:\mathbf{v}_n \in S_i} v_n(2)\right)\sin(\alpha_i).$$

Also,

$$\frac{\partial \varepsilon_i^2}{\partial \alpha_i} = \frac{2}{M_i}\left(\sum_{n:\mathbf{v}_n \in S_i} (v_n(1)\cos(\alpha_i) + v_n(2)\sin(\alpha_i) - \rho_i) \times (-v_n(1)\sin(\alpha_i) + v_n(2)\cos(\alpha_i))\right)$$

If $\partial \varepsilon_i^2 / \partial \rho_i = 0$, we have the equation

$$\frac{1}{M_i}\left(\sum_{n:\mathbf{v}_n \in S_i} (v_n(1)\cos(\alpha_i) + v_n(2)\sin(\alpha_i))(-v_n(1)\sin(\alpha_i) + v_n(2)\cos(\alpha_i))\right) = \frac{1}{M_i}\left(\sum_{n:\mathbf{v}_n \in S_i} \rho_i(-v_n(1)\sin(\alpha_i) + v_n(2)\cos(\alpha_i))\right).$$

The left side of the above equation can be written as $(1/2)\sin(2\alpha_i)(1/M_i)\Sigma_{n:\mathbf{v}_n \in S_i}(v_n^2(2) - v_n^2(1)) + \cos(2\alpha_i)(1/M_i)\Sigma_{n:\mathbf{v}_n \in S_i} v_n(1)v_n(2)$, while the right side can be written as $(((1/M_i)\Sigma_{n:\mathbf{v}_n \in S_i} v_n(1))\cos(\alpha_i) + ((1/M_i)\Sigma_{n:\mathbf{v}_n \in S_i} v_n(2))\sin(\alpha_i))(((1/M_i)\Sigma_{n:\mathbf{v}_n \in S_i} v_n(2))\cos(\alpha_i) - ((1/M_i)\Sigma_{n:\mathbf{v}_n \in S_i} v_n(1))\sin(\alpha_i))$. Let $a_1 = (1/M_i)\Sigma_{n:\mathbf{v}_n \in S_i} v_n(1)$, $b_1 = (1/M_i)\Sigma_{n:\mathbf{v}_n \in S_i} v_n(2)$, $a_2 = (1/M_i)\Sigma_{n:\mathbf{v}_n \in S_i} v_n^2(1)$, $b_2 = (1/M_i)\Sigma_{n:\mathbf{v}_n \in S_i} v_n^2(2)$, and $c = (1/M_i)\Sigma_{n:\mathbf{v}_n \in S_i} v_n(1)v_n(2)$, we have $\tan(2\alpha_i) = 2(c - a_1 b_1)/((b_1^2 - a_1^2) - (b_2 - a_2))$ or $\alpha_i = (1/2)\tan^{-1}(2(a_1 b_1 - c)/((a_1^2 - a_2) - (b_1^2 - b_2))) + \gamma\pi/2$, where γ is either 0 or 1. Also, $\rho_i = a_1\cos(\alpha_i) + b_1\sin(\alpha_i)$. Because we use ρ_i to represent the distance between the point of the origin and the centroid line, the value of ρ_i should always be non-negative. Thus,

$$\rho_i = |a_1\cos(\alpha_i) + b_1\sin(\alpha_i)|. \tag{A1}$$

If $a_1\cos(\alpha_i) + b_1\sin(\alpha_i) < 0$, we then need to add a constant angle $-\pi$ to α_i. The solution for parameter α_i is either $\alpha_i = (1/2)\tan^{-1}(2(a_1 b_1 - c)/((a_1^2 - a_2) - (b_1^2 - b_2))) - \pi(1 - u(a_1\cos(\alpha_i) + b_1\sin(\alpha_i)))$ or

$$\alpha_i = \frac{1}{2}\tan^{-1}\left(\frac{2(a_1 b_1 - c)}{(a_1^2 - a_2) - (b_1^2 - b_2)}\right) + \frac{\pi}{2} - \pi(1 - u(a_1\cos(\alpha_i) + b_1\sin(\alpha_i))), \tag{A2}$$

where the unit step function $u(\tau)$ is equal to 0 for negative value of τ, 1 otherwise.

REFERENCES

[1] L. G. Alexopoulos, G. R. Erickson, and F. Guilak, "A method for quantifying cell size from differential interference contrast images: Validation and application to osmotically stressed chondrocytes", J. Microsc. **205**, 125–135 (2002).

[2] X. Chen, X. Zhou, and S. T. Wong, "Automated segmentation, classification, and tracking of cancer cell nuclei in time-lapse microscopy", IEEE Trans. Biomed. Eng. **53**, 762–766 (2006).

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

[4] H.-S. Wu, J. Barba, and J. Gil, "Iterative thresholding for segmentation of cell images", J. Microsc. **197**, 296–304 (2000).

[5] H.-S. Wu, R. Xu, N. Harpaz, D. Burstein, and J. Gil, "Segmentation of intestinal gland images with iterative region growing", J. Microsc. **220**, 190–204 (2005).

[6] L. D. Biertho, C. Kim, H.-S. Wu, P. Unger, and W. B. Inabnet, "Relationship between sestamibi uptake, parathyroid hormone assay and nuclear morphology in primary hyperparathyroidism", J. Am. Coll. Surg. **199**, 229–233 (2004).

[7] L. Deligdisch, C. Miranda, H.-S. Wu, and J. Gil, "Human papillomavirus-related cervical lesions in adolescents: A histologic and morphometric study", Gynecol. Oncol. **89**, 52–59 (2003).

[8] T. Mattfeldt, H.-W. Gottfried, V. Schmidt, and H. A., Kestler, "Classification of spatial textures in benign and cancerous glandular tissues by stereology and stochastic geometry using artificial neural networks", J. Microsc. **198**, 143–158 (2000).

[9] C. O. D. Solorzano, E. G. Rodriguez, A. Jones, D. Pinkel, J. W., Gray, D. Sudar, and S. J. Lockett, "Segmentation of confocal microscope images of cell nuclei in thick tissue sections", J. Microsc. **193**, 212–226 (1999).

[10] T. Hortobagyi, S. Wise, N. Hunt, N. Cary, V. Djurovic, A. Fegan-Earl, K. Shorrock, D. Rouse, and S. Al-Sarraj, "Traumatic axonal damage in the brain can be detected using β-APP immunohistochemistry within 35 min after head injury to human adults", Neuropathol. Appl. Neurobiol. **33**, 226–237 (2007).

[11] L. D'Este, A. Casini, S. Puglisi-Allegra, S. Cabib, and T. G. Renda, "Comparative immunohistochemical study of the dopaminergic systems in two inbred mouse strains (C57BL/6J and DBA/2J)", J. Chem. Neuroanat. **33**, 67–74 (2007).

[12] A.-L. Mausset-Bonnefont, R. D. Seze, and A. Privat, "Immunohistochemistry as a tool for topographical semi-quantification of neurotransmitters in the brain", Brain Res. Brain Res. Protoc **10**, 148–155 (2003).

[13] W. T. Shaia, S. M. Shapiro, A. J. Heller, D. L. Galiani, A. Sismanis, and R. F. Spencer, "Immunohistochemical localization of calcium-binding proteins in the brainstem vestibular nuclei of the jaundiced Gunn rat", Hear. Res. **173**, 82–90 (2002).

[14] H.-S. Wu and J. Gil, "Linear clustering for segmentation of color microscopic lung cell images", J. Imaging Sci. Technol. **47**, 161–170 (2003).

[15] R. J. Schalkoff, *Digital Image Processing and Computer Vision* (Wiley, New York, 1989).

[16] A. C. Ruifrok and D. A. Johnston, "Quantification of histochemical staining by color deconvolution", Anal Quant Cytol. Histol. **23**, 291–299 (2001).