CELL TRACKING AND SEGMENTATION IN ELECTRON MICROSCOPY IMAGES USING GRAPH CUTS

Huei-Fang Yang and Yoonsuck Choe

Department of Computer Science and Engineering
Texas A&M University
3112 TAMU, College Station, TX 77843-3112

ABSTRACT

Understanding neural connectivity and structures in the brain requires detailed 3D anatomical models, and such an understanding is essential to the study of the nervous system. However, the reconstruction of 3D models from a large set of dense nanoscale medical images is very challenging, due to the imperfections in staining and noise in the imaging process. Manual segmentation in 2D followed by tracking the 2D contours through cross-sections to build 3D structures can be a solution, but it is impractical. In this paper, we propose an automated tracking and segmentation framework to extract 2D contours and to trace them through the z direction. The segmentation is posed as an energy minimization problem and solved via graph cuts. The energy function to be minimized contains a regional term and a boundary term. The regional term is defined over the flux of the gradient vector fields and the distance function. Our main idea is that the distance function should carry the information of the segmentation from the previous image based on the assumption that successive images have a similar segmentation. The boundary term is defined over the gray-scale intensity of the image. Experiments were conducted on nanoscale image sequences from the Serial Block Face Scanning Electron Microscope (SBF-SEM). The results show that our method can successfully track and segment densely packed cells in EM image stacks.

***Index Terms*—** graph cuts, electron microscopy, image segmentation

1. INTRODUCTION

Understanding neural connectivity and structures in the brain requires detailed 3D anatomical models. To provide the data for detailed 3D connectivity of neuronal circuits, Denk and Horstmann [1] recently developed Serial Block Face Scanning Electron Microscopy (SBF-SEM) to generate nanoscale images to make the imaging of 3D microstructures possible. SBF-SEM data are a stack of 2D nanoscale medical images with a resolution on the order of tens of nanometers. The sectioning thickness is around $30nm$ and the lateral resolution can achieve as high as 10-20nm/pixel. The high image resolution allows researchers to identify small organelles, even to trace axons and to identify synapses. Interstitial staining for SBF-SEM images highlight cell boundaries so that cells (foreground) are in brighter gray-scale intensity and non-cells (background) are in darker gray-scale intensity.

The goal of the 3D reconstruction of a SBF-SEM image stack is to segment regions in each 2D image first and then to find the region correspondences between adjacent images. Segmentation of SBF-SEM images amounts to delineating cell boundaries. The challenges of identifying cell boundaries come from inevitable staining noise and weak or missing boundaries between cells that are located very close. In this paper, we propose a tracking and segmentation framework to overcome the boundary ambiguity problem. In our segmentation framework, an assumption is made that the region appearances between adjacent images are similar, so the shape information from the previous segmentation can facilitate the determination of correct region boundaries in the current image. Segmentation is posed as a Maximum A Posteriori estimation of a Markov Random Field (MAP-MRF) energy function minimization problem. The energy function to be minimized has a regional term and a boundary term. The regional term is defined over the flux of the image gradient vector fields from the current image and the shape information from the previous image which is formulated as the distance function. The boundary term is defined over the image gray-scale intensity. Finally, graph cut technique was used to obtain the globally optimal solution.

In related work, [2] developed an active contour based method to track and segment axons. However, their method failed when the region boundaries are missing or blurred. [3] proposed a semi-automated level-set segmentation method allowing the user to adjust the segmentation results. However, the computation of level-set method is expensive and also the solution can sometimes get stuck in local minima. Our segmentation method using graph cuts to minimize the energy function guarantees to obtain the globally optimal solution.

This work was supported in part by NIH/NINDS grant #1R01-NS54252. We would like to thank Stephen J. Smith (Stanford) for the SBF-SEM data.

The remainder of this paper is organized as follows. In section 2, a brief review of graph cut segmentation technique is given. In section 3, a detailed description is presented as to how to determine the flux of the gradient vector fields and distance function in the regional term of the energy function. Section 4 gives the definition of boundary term in the energy function. Section 5 shows the experimental results on SBF-SEM data, followed by a conclusion and future work in section 6.

2. GRAPH CUT SEGMENTATION

Graph cut technique has recently been applied successfully to image segmentation since Boykov and Jolly [4] introduced it into the computer vision field. Image segmentation can be considered as a binary labeling problem. Assume that an image is represented as a graph $G = \langle V, E \rangle$ with a set of vertices (nodes) V representing pixels or image regions and a set of edges E connecting the nodes. Each edge is associated with a nonnegative weight. The binary labeling problem is to assign each node i with a unique label x_i, that is, $x_i \in \{0\,(\text{background}), 1\,(\text{foreground})\}$ such that $X = \{x_i\}$ minimizes the following energy function:

$$E(X) = \lambda \cdot R(X) + B(X) \tag{1}$$

where λ is a positive parameter for adjusting the relative weighting between R and B, and

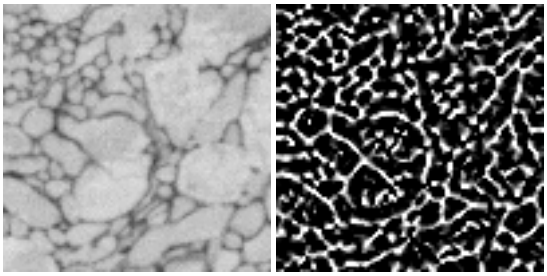
$$R(X) = \sum_{i \in V} E_i(x_i) \ \text{(regional term)} \tag{2}$$

$$\begin{aligned} B(X) &= \textstyle\sum_{\{i,j\} \in N_i} E(x_i, x_j) \\ &= \textstyle\sum_{\{i,j\} \in N_i} w_{i,j} \cdot (1 - \delta(x_i, x_j)) \ \text{(boundary term)} \end{aligned} \tag{3}$$

where N_i is the neighbors of node i, $E(x_i, x_j)$ denotes the cost definition function on nodes i and j when their assigned labels are x_i and x_j, respectively, $w_{i,j}$ is the weight between nodes i and j, and the indicator function $\delta(x_i, x_j)$ is defined as:

$$\delta(x_i, x_j) = \begin{cases} 0 & \text{if } x_i \neq x_j \\ 1 & \text{if } x_i = x_j \end{cases}$$

$E_i(x_i)$ in the regional term indicates the cost when a label x_i is assigned to a node i. $E(x_i, x_j)$ in the boundary term captures the cost when nodes i and j are assigned different labels, i.e., there is a discontinuity between nodes i and j. The binary labeling problem described above is posed as a Maximum A Posteriori estimation of a Markov Random Field (MAP-MRF) and can be minimized by graph cuts. To minimize equation (1) by graph cuts, two additional nodes, s and t, are introduced. s and t denote a foreground terminal and a background terminal, respectively. All nodes $i \in V$ are linked to the terminals s and t with weight $w_{s,i}$ and $w_{i,t}$. For more information as to how to solve the minimization problem by graph cuts, please refer to [5].



(a) SBF-SEM image (b) associated flux of (a)

Fig. 1. Flux of the gradient vector fields of slice 1.

3. REGIONAL TERM

A regional term defines the cost that a node is assigned a label x_i, that is, the corresponding cost of assigning node i to the foreground and the background. In our segmentation framework, a regional term consists of two parts, the flux of the gradient vector fields and the distance function carrying the shape information from the segmentation of the previous image in the image stack.

3.1. Flux

Flux has recently been introduced by Vasilevskiy and Siddiqi [6] into image analysis and computer vision. They incorporated the flux into level-set method to segment blood vessel images. After that, the flux has also been integrated into graph cuts [7] [8] to improve the segmentation accuracy. The introduction of flux into graph cuts can reduce the discretization artifacts which is the major shortcoming in graph cuts [7]. According to [7], the flux of a given vector field v through a given continuous hypersurface S is

$$F(S) = \int_S \langle N, v \rangle \, dS \tag{4}$$

where $\langle , \rangle$ is the Euclidean dot product, N are unit normals to surface element dS consistent with a given orientation. Inward and outward are two possible orientations that can be assigned to S. In our experiments, the flux is computed in 2D. v is defined as the gradient vector field of the smoothed image and N is the outward normal at each point on the contour. The flux of a point i is obtained by summing from the eight neighbors of point i the values computed by the dot product of the gradient vector field and the unit outward normal at i of the unit disc centered at i [9]. Fig. 1(b) shows the flux of gradient vector fields of fig. 1(a). The foreground object has negative flux (dark) whereas the background has positive flux (bright).

3.2. Distance Function

The assumption is that the region appearances between adjacent images should be similar, thus the shape information from the previous segmentation can be formulated as a constraint in the energy function. The constraint enables graph cuts to determine the correct boundaries if missing or blurred boundaries are encountered. The region shape is represented by a distance function. Let O_{t-1} denote the object in image $t-1$. The distance function $D(i)$ of a pixel i in image t is

$$D(i) = \begin{cases} \|i - o_i\| & i \text{ is outside } O_{t-1} \\ 0 & i \text{ is inside } O_{t-1} \end{cases} \quad (5)$$

where $\|i - o_i\|$ represents the Euclidean distance from i to the nearest object pixel $o_i \in O_{t-1}$. The distance function penalizes pixels outside the previously segmented objects but pixels inside the previously segmented objects get no penalty. In other words, the larger the distances between the pixels outside the previously segmented objects and previously segmented objects' boundaries, the lower the possibility of those pixels belonging to the foreground. One may argue that the choice of the above distance function is not symmetric for the foreground and the background. Another symmetric distance function to the foreground and the background was also implemented and tested, however, the asymmetric and symmetric distance functions both yielded similar segmentation results.

3.3. Incorporating Flux and Distance Function into Regional Term

Combining the flux of gradient vector fields of the current image with distance function from the previous segmentation yields a new regional term. Inspired by [8], the edge weights between node i and terminals s and t are assigned as:

$$\begin{aligned} w_{s,i} &= -\min(0, F(i)) \\ w_{i,t} &= \max(0, F(i)) + \alpha D(i) \end{aligned} \quad (6)$$

where $F(i)$ denotes the flux at point i, and α is a positive parameter to adjust the relative importance of the distance function. In our experiments, the value of α was set to 0.2.

4. BOUNDARY TERM

In SBF-SEM images, the foreground and background can be discriminated by their gray-scale intensities. Compared to the background, the foreground objects usually have higher intensity values. Boundaries can thus be determined if the intensity differences between pixels are large. To capture the boundary discontinuity between pixels, the weight between node i and its neighbor j is defined as:

$$w_{i,j} = \exp\left(-\frac{(I_i - I_j)^2}{2\sigma^2}\right) \cdot \frac{1}{\|i - j\|} \quad (7)$$

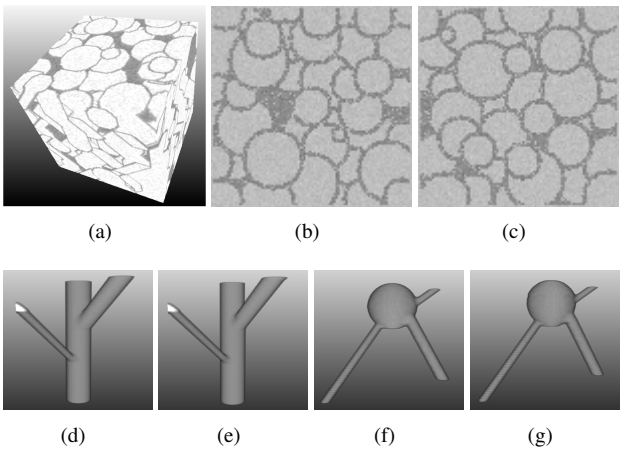
where I_i and I_j are pixel intensities ranging from 0 to 255, $\|i - j\|$ is the Euclidean distance between i and j, and σ is a positive parameter set to 30. The above equation penalizes a lot for edges with similar gray-scale intensities between nodes while it penalizes less for those with larger gray-scale differences between nodes. That is, the edge weights are small at the boundaries, where a cut is more likely to occur. Note that an 8-neighborhood system is used in our experiments.

5. EXPERIMENTAL RESULTS

The proposed tracking and segmentation method was performed both on synthetic data and on an SBF-SEM image stack from the optic tectum of larval zebrafish. Only for the very first image in the stack, the user manually delineates the boundaries of the regions for the algorithm to track and to segment. The manually segmented regions serve as the distance function in segmenting regions in the subsequent image. After that, the shape information is obtained from the segmentation results generated by the proposed method. The synthetic data consists of two image stacks with size $100 \times 100 \times 100$. Noise was added to each image slice to simulate the noise in SBF-SEM images. Fig. 2(a) shows one of the synthetic image stacks and fig. 2(b) and fig. 2(c) show the selected noisy image slices from the two synthetic image stacks. Fig. 2(e) and fig. 2(g) show the 3D reconstruction from the synthetic image slices. The ground truth of the synthetic data is shown in fig. 2(d) and fig. 2(f). We can see that the reconstruction is almost identical to the ground truth with minor differences. Experiments on SBF-SEM data were conducted on one image stack ($631 \times 539 \times 560$), on different parts (sub-volumes) of it. Fig. 3 demonstrates the tracked and segmented 2D contours on selected image slices of the SBF-SEM data, where the user manually delineates one region boundary on the first image (the user can delineate more regions if desired), and the algorithm automatically tracks and segments the 2D contour in the subsequent images. Fig. 4(a) shows the SBF-SEM image stack. Fig. 4(b) and fig.4(c) show the reconstructed 3D structures of tracked and segmented 2D contours, where part of a neuron can be seen.

6. CONCLUSION

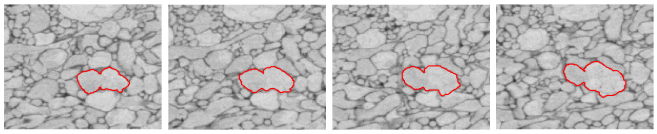
We presented a novel tracking and segmentation framework to extract 2D contour boundaries from SBF-SEM image stacks. Our main idea was that the segmentations between adjacent images in the image stack should be similar, thus the segmentations from the previous image can be used as the distance function to delineate the region boundaries in the current image. Graph cuts were used to obtain a globally optimal solution of an energy function which consists of the flux of the gradient vector fields, the image gray-scale intensity, and the distance function. The advantage of incorporating the distance function into an energy function is that it can solve



(a) (b) (c)

(d) (e) (f) (g)

Fig. 2. Synthetic image stack, noisy images and extracted 3D volumes of synthetic data. (a) is one of the synthetic image stacks. (b) and (c) show the selected noisy image slices from the two synthetic image stacks. (d) and (f) are the ground truth of the synthetic data. (e) and (g) are the reconstructed 3D volumes of the $100 \times 100 \times 100$ synthetic image stacks. Our results are very close (almost identical) to the ground truth, with only very minor differences.



(a) slice 1 (b) slice 5 (c) slice 10 (d) slice 15

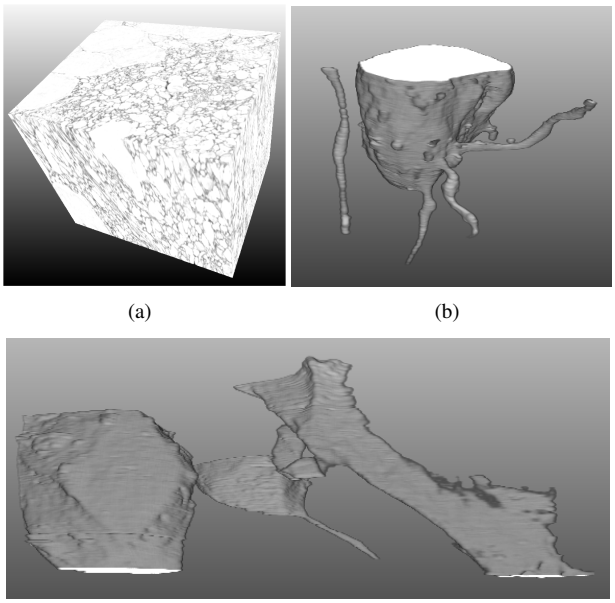
Fig. 3. Tracked and segmented 2D contours on selected image slices. The user manually delineates the boundary of the region for tracking and segmentation in the first image. The algorithm tracks and segments the selected region in the subsequent images.

the boundary ambiguity problem occurring in densely packed EM images. Future work includes applying the method to larger 3D volumes and seeking a quantitative and qualitative validation method on the segmentation results obtained from SBF-SEM data set which currently lacks the ground truth.

7. REFERENCES

[1] W. Denk and H. Horstmann, "Serial block-face scanning electron microscopy to reconstruct three-dimensional tissue nanostructure," *PLoS Biology*, vol. 2, no. 11, pp. e329, 2004.

[2] E. Jurrus, T. Tasdizen, P. Koshevoy, P. T. Fletcher, M. Hardy, C. Chien, W. Denk, and R. Whitaker, "Axon tracking in serial block-free scanning electron microscopy," in *MIAAB*, 2006, pp. 114–119.



(a) (b)

(c)

Fig. 4. SBF-SEM image stack and partly reconstructed 3D structures from an SBF-SEM image stack. (a) is the SBF-SEM image stack. (b) and (c) are the partly reconstructed 3D structures, where part of a neuron can be seen in (b).

[3] J. H. Macke, N. Maack, R. Gupta, W. Denk, B. Schölkopf, and A. Borst, "Contour-propagation algorithms for semi-automated reconstruction of neural processes," *Journal of Neuroscience Methods*, vol. 167, no. 2, pp. 349–357, 2008.

[4] Y. Boykov and M. Jolly, "Interactive graph cuts for optimal boundary and region segmentation of objects in N-D images," in *ICCV*, 2001, pp. 105–112.

[5] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 11, pp. 1222–1239, 2001.

[6] A. Vasilevskiy and K. Siddiqi, "Flux maximizing geometric flows," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 12, pp. 1565–1578, 2002.

[7] V. Kolmogorov and Y. Boykov, "What metrics can be approximated by geo-cuts, or global optimization of length/area and flux," in *ICCV*, 2005, pp. 564–571.

[8] N. Vu and B. S. Manjunath, "Graph cut segmentation of neuronal structures from transmission electron micrographs," in *ICIP*. 2008, pp. 725–728, IEEE.

[9] P. Dimitrov, C. Phillips, and K. Siddiqi, "Robust and efficient skeletal graphs," in *CVPR*, 2000, pp. 1417–1423.