31st Annual International Conference of the IEEE EMBS
Minneapolis, Minnesota, USA, September 2-6, 2009

Neuron Branch Detection And Description Using Random Walk

Hee Chang Kim and Auguste Genovesio

Abstract— **The morphological studies of neuron structures are of great interests for biologists. However, manually detecting dendrites structures is very labor intensive, therefore unfeasible in studies that involve a large number of images. In this paper, we propose an automated neuron detection and description method. The proposed method uses ratios of probability maps from random walk sessions to detect initial seed-points and minimal cost path integrals with Delaunay triangulations.**

I. INTRODUCTION

Neurons are the basic building blocks of the nervous system. Each part of a neuron plays a role in the communication of information throughout the body. Therefore, qualitative morphological studies of neurons are of great interest. However, there are numerous challenges due to their structural complexities and sometimes poor image qualities. To this end, various methods were proposed in the literature such as ones based on direct exploratory tracing [1], [2] which stems from [3], and probabilistic segmentation method [4] to name a few. As a prerequisite, [1] and [2] searches for seed-points that ensure proper start locations and directions of tracing. In particular, two-step process, which includes line searches and filtering out points on the background, is used. A modified version of the algorithm, presented by [3], is used to begin neurite tracing until some conditions of stopping criteria are reached.

In this paper, we propose a modified probabilistic segmentation proposed by [4] to obtain reliable seed-points and applying minimum cost approach [5] in neuron description. Probability maps of two different models of random walk sessions, intensity-biased and unbiased, are computed and used to obtain and separate sets of seed-points on a neuron cell body and its branches. Then minimum cost approach is applied along with Delaunay triangulation in an attempt to describe branches of the neuron.

In section II, a brief introduction to random walk model used in our method and its application is presented. In section III, we explain the motivations behind our proposed method and show the framework of the method in each subsection. Then we conclude in section IV.

This work was supported by Institut Pasteur Korea

Hee Chang Kim is with Systemes intelligents de perception Universite Rene Descartes Paris 5, 45, rue des Saints-Peres, 75006 Paris caleb007@ip-korea.org

Auguste Genovesio is with Image Mining Group, Institut Pasteur Korea, Sampyeong-dong Bundang-gu, Gyeonggi-do, Seongnam-si, Korea pmisra@cs.wright.edu

II. BACKGROUND

We consider a random walk session starting from a given point on an image. Suppose that a random walk begins from a position somewhere in the cell body (soma) of a neuron, obtained by a method described in section III-B. For each step, we determine the directions that the walker takes by calculating the normalized surrounding intensities of the position that the walker is currently in. Since we expect the walker's excursions to stay inside the neuron cell body and its branches, and under the assumption that there are sharp drops in intensity along the cell membrane, we assign higher probability of the direction taken by the walker to the neighboring pixels that has similar intensities. Then after a sufficient number of steps of excursion, we obtain which pixels are more likely to have been visited by the walker whose directions of steps are influenced by local intensities.

In the beginning stage of our study, a simulated model of such random walk sessions was built. However there were an ambiguity on how to keep a visit count and unclear correlation between the visit count and the probability distribution of the random walk. To be specific, we asked a question of whether a visit count should be incremented each time the walker visits a certain position during the excursion or after an excursion is finished. When we took the former idea, we got visit records that resembled geometric distributions centered at the starting position of the excursions. And when the latter was taken, the numbers were too sporadic around the center that we could not consider the result reliable. This led us to adopt a finite Markov chain model that computes probability distributions of a random walk position after a certain number of steps over the image.

In order to build an iterative Markov chain model, we first form a transition matrix that is column normalized as follows. Let $I(i,j)$ be an image with width m and height n, and let $\mathbf{P}$ be the transition matrix with $[m*n]$ rows and $[m*n]$ columns. Each column of the matrix corresponds to a pixel of the image $I(i,j)$ such that $(x_0+y_0*m)^{th}$ column represents a pixel (x_0,y_0). The elements of the column is calculated by first normalizing the gradient intensities of the neighbors of (x_0,y_0) then assigning the obtained values to appropriate rows of the $(x_0+y_0*m)^{th}$ column. In normalizing the gradient intensities of the neighbors, we take the typical Gaussian weighting functions commonly used in the literature [6]. Let (x_0,y_0) be the intensity at current position of the random walk session and (x_k,y_k) one of its neighbors . Then the element in the corresponding

position of the column is

$$\mathbf{P}\left(x_k + y_k * m, x_0 + y_0 * m\right) = e^{-((x_0,y_0)-(x_k,y_k))^2}. \quad (1)$$

As we can see, $\mathbf{P}$ is a column normalized sparse matrix with only 8 non-zero elements in each column. With the transition matrix formed, we need to define two column vectors $\mathbf{s}$ and $\mathbf{x}_{(x_0,y_0)}$ such that all elements of them are zeroes except at the starting position of the random walk which is 1. We also define a value c which is a restart probability that determines the range of area covered by the random walk. Finally, we can iteratively obtain a probability mask $\mathbf{x}_{(x_0,y_0)}$:

$$x_{(x_0,y_0)} := (1-c)\, Px_{(x_0,y_0)} + cs, c \in [0,1]\,. \quad (2)$$

The authors in [7] show that Eq. (2) converges to a stationary probability distribution. If c is close to 1, then the probability mask covers small area around the starting position and if it is close to zero, it covers a large area on the image.

III. METHOD

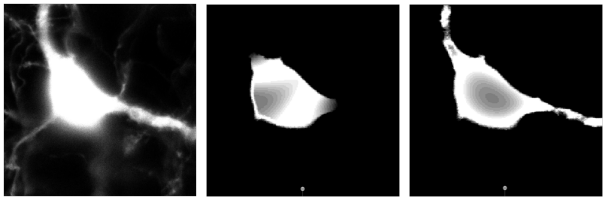
In this section, we first discuss the motivation of our method and compare our results with the ones that were presented by [4]. We then discuss further how our method can be applied to neuron description. We end by proposing a Delaunay Triangulation combined with minimum cost path algorithm in order to detect neurons.

A. Motivation

From the method that was presented in section II, we get, for each pixel on an image, a probability of whether a pixel belongs to the cell that was segmented by the algorithm. This is very useful when there are more than one cell on the image since separate random walk sessions which start from different starting positions give different probability of the pixel belonging to corresponding cells.

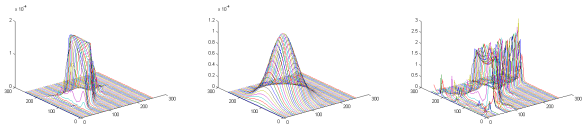
Since our aim is to identify and describe neuron structures, we develop the method further. In [4], the elements of the transition matrix $\mathbf{P}$ is calculated by the normalized intensities of the neighbor pixels with Eq. (1). Also, because we let the random walk excursions to have higher probability of taking directions to neighboring pixels with similar intensity values, we are assured that the centerline positions of the branches have higher probability values compared to the ones near the edges.

However, the resulting probability map Fig. 1(b) shows smooth downward or upward trends of probabilities whence we can not easily locate where the peaks are. Therefore, in addition to transition matrix $\mathbf{P}$, we also form another transition matrix $\mathbf{Q}$. The transition matrix $\mathbf{Q}$ represents a classical random walk that has equal chance to go in any direction for each step. Then the resulting probability mask shows how an unbiased random walk behaves from the same starting position.



(a) (b) (c)

Fig. 1. Neuron segmentation results. (a) Original image. (b) Intensity based random walk. (c) Probability quotient.



(a) (b) (c)

Fig. 2. 3D plot of neuron segmentation results. (a) Segmentation result from [4]. (b) Unbiased random walk result. (c) Probability Mask $z_{pq}(i,j)$.

Let $\mathbf{x}_p$ and $\mathbf{x}_q$ be the resulting probability masks from Eq. (2) with transition matrices $\mathbf{P}$ and $\mathbf{Q}$ respectively. Then $x_p(i,j)$ is the probability that gradient intensity biased random walk ends up at (i,j) after a certain number of steps and $\mathbf{x}_q(i,j)$ is the probability that unbiased random walk finishes at (i,j) after the same number of steps. Thus,

$$z_{pq}(i,j) = x_p(i,j)/x_q(i,j)\,, \quad (3)$$

gives probabilistic values for an intensity biased random walk to be in the position compared to the unbiased one after a certain number of steps taken. Fig. 1(c) shows the resulting images from Eq. (3).

Notice how we can obtain more intense pixel values at the branches on Fig. 1(c). In Fig. 2, 3D plots of probability masks from intensity biased (a) and unbiased (b) are displayed. Also, the third 3D plot (c) shows clearly the peaks along the branches of the neuron. Hence, we obtain locations of peaks inside the neuron without applying thresholds in order to discount the false positive positions outside the neuron which we must do if we used methods in [2].

B. Soma detection

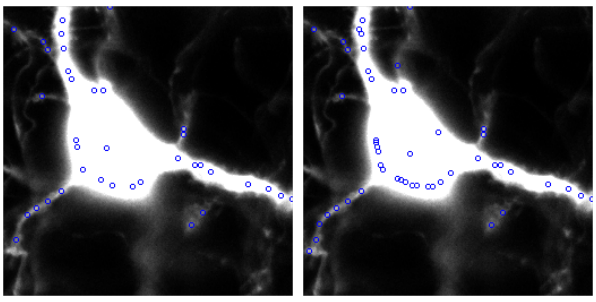
As mentioned in section II, soma (cell bodies of neurons) must first be found in order to run the processes shown in section III-A. As proposed in [2], we use a simple gray-scale erosion with an assumption that the size of soma is bigger in pixel than branches. Since the width of the soma is much larger than the width of dendrites, we can easily find a window size with which we run a gray-scale erosion. Then we look for local maxima over the whole image which yields positions near the center of soma. This method is effective in detecting the location of soma, and we find that little

variations in locations of random walk starting positions inside soma does not alter the accuracy of the results.

C. Seed-points acquisition and classfication

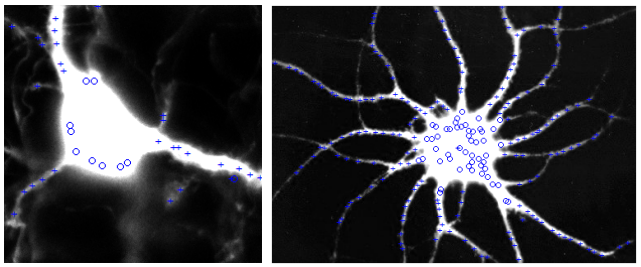
In section III-A, we showed that by using Eq. (3) we locate the peaks of probabilistic values (referred to as salient points hereafter) over the neuron. A simple local maxima algorithm is run to locate positions of the peaks. We may obtain higher number of salient points as we adjust the size of the window used in the local maxima algorithm. Fig. 3 shows the locations of the salient points. Fig. 3(a) displays the salient points found with a window size 5 and (b) with a window size 3. Notice how the locations of the salient points are nearly at the centerlines of the branches and that there are numerous peaks found along the edge of the soma. The two step process of first locating soma position and then running random walk sessions detect salient points that are inside neuron. In fact, it dramatically reduces false positives such that we do not need to apply a threshold in order to get rid of salient points that lie outside the neuron.

Now we must distinguish the salient points on the soma from the ones on branches. To this end, we collect all probability values at salient point locations from the probability mask x_p (Fig. 2(a)). Since we are assured that the intensities on soma are higher than intensities on branches, which means that probabilities of the intensity biased random walk on soma are higher, we apply K-means algorithm on collected probability values to separate salient points that are on soma from the ones on branches. Fig. 4 shows two resulting images after running the procedures in previous sections and classifying obtained peak locations with K-means algorithm into two groups. A salient point is marked with a circle if a point is on the soma or marked with a "+" otherwise.



(a) (b)

Fig. 3. Locations of salient points obtained from finding local maxima on probability mask $z_{pq}(i,j)$. (a) window size 5. (b) window size 3.



(a) (b)

Fig. 4. Classification of salient points into soma points and branch points.

D. Delaunay triangulation and minimum cost path

From section III-B and III-C, we obtain the locations of the center of the soma and salient points that are on the branches of the neuron. It remains to find minimal paths between salient points and the center of the soma with a condition that a salient point is connected to either the center point or the nearest salient point. A fast way to construct all possible paths that satisfy the condition is to run a Delaunay triangulation with all the salient points and the center point. Fig. 5(a) displays the image after the triangulation is done. Since the triangulation includes all possible paths of salient points to the center point, we must eliminate edges that are not representative of lines on the image. In doing so, we adopt a cost function that was introduced in [5]. From the second order directional derivatives on the image at a certain observation scale σ, we calculate the Hessian matrix

$$H = \begin{pmatrix} I_{xx} & I_{xy} \\ I_{yx} & I_{yy} \end{pmatrix}$$

and choose $\lambda_{\max}$ such that $\lambda_{\max} = \max(|\lambda_1|, |\lambda_2|)$ where λ_i is an eigenvalue of the Hessian matrix. Define a quantitative measure of line contrast $R_\sigma(x,y)$ by:

$$R_\sigma(x,y) = \sigma^2 |\lambda_{\max}| \frac{1}{b^\sigma}, \tag{4}$$

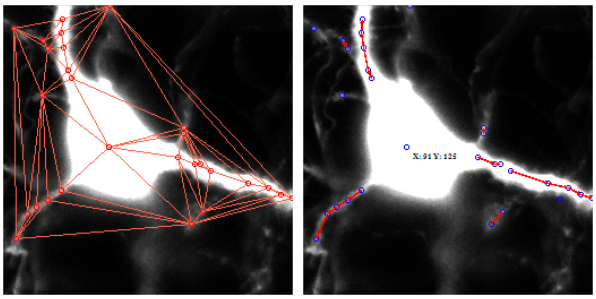
where the line brightness b is given by

$$b^\sigma = \begin{cases} I^\sigma_{(x,y)} & \text{if } \lambda_{\max} \leq 0 \\ W - I^\sigma_{(x,y)} & \text{otherwise} \end{cases} \tag{5}$$

where $W = \max(I(x,y))$. We take $W = \max(I(x,y))$ since the lines formed by the branches of the neuron are bright lines against a low intensity background. From this quantitative figure R_σ, we determine to what degree a point on an image belongs to a line.

Assume that S_1 and S_2 are two end points of an edge from the triangulation. Using Brensenham's algorithm, we obtain all points that belong to a path between S_1 and S_2. Let $r_\sigma(x,y)$ be a cost function given by:

$$r_\sigma(x,y) = \frac{\epsilon}{\epsilon + R_\sigma(x,y)}, \tag{6}$$



(a) (b)

Fig. 5. Delaunay triangulations with obtained salient points. (a) After Delaunay triangulations. (b) After edge elimination by minimum cost path algorithm.

and its path integral

$$c\left(S_1, S_2\right) = \int_{S_1}^{S_2} r_\sigma\left(x\left(p\right), y\left(p\right)\right) dp, \tag{7}$$

where $\left(x\left(p\right), y\left(p\right)\right)$ is the position of the point on the path and fixed ϵ that represents a trade-off between following the maximum line contrast and the shortest route. The path integral (7) yields integrated cost over the path that is given by the edges of the triangulation at a given scale σ. Then we calculate the cost path integrals for all edges from the triangulation and eliminate the edges with path integral values higher than a number that is given by the user. The preliminary result of the minimal cost path analysis is shown in Fig. 5(b) with $\sigma = 2$ and $\epsilon = 0.001$. We can see that the most of the edges that do not represent lines were eliminated and three major dendrites branching out of the soma are kept.

IV. CONCLUSION

In this paper, we presented a method that uses probability distributions from random walk sessions to detect reliable seed-points and minimum cost path analysis to obtained seed-points for extracting, constructing, and describing major branches that neurons have. The main contributions of this paper are first the method's utilization of two different types of random walk sessions in order to extract salient points as opposed to just one in [4] for segmentation and second its ability to extract more reliable salient points thereby eliminating the need to threshold the falsely detected salient points as done in [2]. The preliminary results are encouraging and we plan to extent this work by refining the minimum cost path analysis to build better initial sketches of the neuron structure and finding the actual minimum paths between the connected salient points from the initial sketches.

REFERENCES

[1] K.A. Al-kofahi, S.Lasek, D.H. Szarowski, C.J. Pace, G. Nagy, (Senior Member), J. N. Turner, and B. Roysam, "Rapid automated three-dimensional tracing of neurons from confocal image stacks," *IEEE Transactions on Information Technology in Biomedicine*, vol. 6, pp. 171–187, 2002.

[2] S.K. Guduru, S.V. Narasimhan, S.T. Birchfield, and B.Z. Gao, "Analysis of neurite outgrowth for a laser patterned neuronal culture," *Neural Engineering, 2005. Conference Proceedings. 2nd International IEEE EMBS Conference on*, pp. 336–339, March 2005.

[3] A. Can, H. Shen, J.N. Turner, H.L. Tanenbaum, and B. Roysam, "Rapid automated tracing and feature extraction from live high-resolution retinal fundus images using direct exploratory algorithms," *IEEE Trans. on Info. Tech. for Biomedicine*, vol. 3, pp. 125–138, 1999.

[4] V. Ljosa and A.K. Singh, "Probabilistic segmentation and analysis of horizontal cells," in *ICDM '06: Proceedings of the Sixth International Conference on Data Mining*, Washington, DC, USA, 2006, pp. 980–985, IEEE Computer Society.

[5] J. Geusebroek, A.W. M. Smeulders, and H. Geerts, "A minimum cost approach for segmenting networks of lines," *International Journal of Computer Vision*, vol. 43, pp. 99–111, 2001.

[6] L. Grady, "Random walks for image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 11, pp. 1768–1783, 2006.

[7] O. Camoglu, T. Can, and A. K. Singh, "Integrating multi-attribute similarity networks for robust representation of the protein space," *Bioinformatics*, vol. 22, no. 13, pp. 1585–1592, 2006, Cited By (since 1996): 1.