EUSIPCO 2013 1569743883

CHASING THE NEUROME: SEGMENTATION AND COMPARISON OF NEURONS

Suvadip Mukherjee[1], Barry Condron[2] and Scott T. Acton[1]

[1]Charles L. Brown Department of Electrical and Computer Engineering
[2]Department of Biology
University of Virginia, Charlottesville, USA.

ABSTRACT

The neuronal content of an organism, the individual morphology of each neuron and the variability of these components constitute the atlas of the neurome. The description of such an atlas will be critical in determining the complex neural system of a given organism, eventually providing clues to how animals think and function. As the organisms under investigation scale from the worm to the human, the number of neurons scale from tens to millions. Image analysis is a key ingredient in tackling a neurome for complex organisms. Specifically, two major problems stand between the state of the art and successful automation. First, neurons must be segmented from microscopy, yielding a morphological description of each individual neuron. Second, these neurons must be matched to prototypes and classified by function. This report describes current progress on these two fronts, revealing encouraging progress as demonstrated on the fruit fly *Drosophila*.

Index Terms— confocal microscopy, image analysis, segmentation, neuron tracing.

1. INTRODUCTION

The *neurome* is to the neurons of an organism as the *genome* project is to genetic content. Along with collaborators in biology, we are developing tools that aid in automating the building of such a neuroma, concentrating of that of the fruit fly *Drosophila*. With around 20,000 neurons in the Ventral Nervous Cord (VNC) of an adult fruit fly *Drosophila* and more than 10^{11} neurons in the adult human brain, the task of developing a complete neural atlas for such organisms is both relevant and daunting. While the neurome of the relatively simpler *C. elegans* is well understood, mapping the neural anatomy of more complicated organisms, such as the neuroma of the fruit fly, mouse and human, is still an unsolved challenge.

Equipped with the state of the art technology to facilitate 3-D imaging of the neurons, researchers are focusing attention on segmenting/tracing neurons from 3-D images. Essentially, two unsolved problems pose challenge in developing the neural anatomy of a species. The first challenge is to be able to trace/segment the neurons from a 3-D stack of microscopy images. Confocal microscopy is widely used in the biological community to image the neurons. The low SNR and discontinuous structure of the images provide hurdles, especially in achieving complete automation of the tracing process. The second challenge, receiving relatively less attention in the research community, is the development of algorithms that categorize pre-segmented neurons into different functional classes based on morphological structure.

In this paper, we give an overview of the recent developments in the field and in our laboratory with an emphasis on the technical challenges and unresolved issues pertaining to both the tracing and matching problems.

2. NEURON SEGMENTAION

Confocal microscopy is a widely used technique to obtain 3-D images of neurons. Despite its benefits, the images obtained using confocal microscopy suffer from significant noise and poor contrast. Figure 1 shows sample images obtained using confocal microscopy. Both the images are plagued by background clutter, low resolution and poor contrast between the foreground and background.

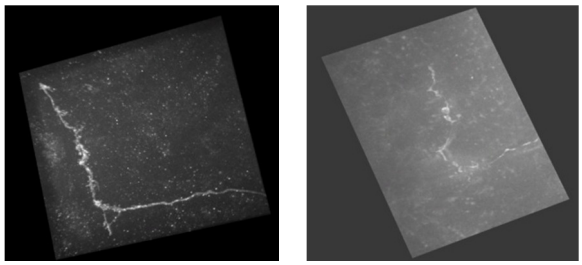


Figure 1: 3-D confocal microscopy images.

An efficient automated segmentation algorithm should be able to effectively trace the neuron from the noisy image, with minimal user interaction, to generate the neuronal tree. The popular tracing methods may be broadly categorized into two classes. The first set of algorithms depends on a set of pre-selected seed-points to initialize the tracing process. The seed points may be selected interactively via user-input, or may be created automatically based on finding points of

interest in the image. Seed point based methods are essentially local processing techniques, since they analyze the regions local to the selected seed-points. While manual seed-selection ensures better accuracy, the inconvenience in terms of time and human effort calls for automated techniques. The robustness of these seed-based techniques is in general affected by the signal intensity and presence of noise. Discontinuity in the neurite structure also creates bottlenecks, resulting in incomplete or erroneous segmentation.

However, seed based approaches are computationally efficient, thus encouraging researchers to develop techniques to overcome the above mentioned difficulties. The widely used neuron tracing tool Vaa3d [1] requires manually chosen seed points for tracing purpose. Although this method requires human supervision, Vaa3d serves as an efficient tool to generate ground truth for neuron segmentation. Recent addition in Vaa3d includes an all path pruning approach [2], which is capable of automatic seed point selection based on an initial over segmentation technique, followed by a specialized pruning process to remove the over-estimated branches. The connectivity between the broken fragments is established by using a graph deformation model, which uses a shortest path approach to connect the broken components [3]. The efficiency of these methods is limited by the choice of some heuristically chosen parameters and the amount of clutter present in the image. Also, the traced path is not guaranteed to lie along the neuron centerline, requiring further manual edits to rectify the error.

Wang et.al. proposed an active contour based methodology for neuron tracing [4]. In this technique the 3-D image is preprocessed to enhance the neuronal structures followed by automated selection of seed points. An open ended snake is initialized to trace each neuronal branch by minimizing an optimization function that forces the active contour to evolve along the neuron centerline. The authors use GVF based technique [5] to develop a medialness measure which may be sensitive to noise. However, this active contour based method is an efficient tool for neuron segmentation, especially in handling datasets with less complicated branching patterns. Typically, the seed-based approaches are computationally efficient, but require special processing techniques to handle branching and end point detection.

The second category of neuron segmentation algorithms is a global approach as opposed to the local seed based tracing methods. These techniques generally consist of four steps: image preprocessing, an initial segmentation, graph based connectivity analysis of the possibly disjoint components and an optional pruning step. Tree2Tree [6] and Tree2Tree-2 [7] are two global neuron segmentation techniques. A brief overview of the Tree2Tree-2 is shown in Figure 2.

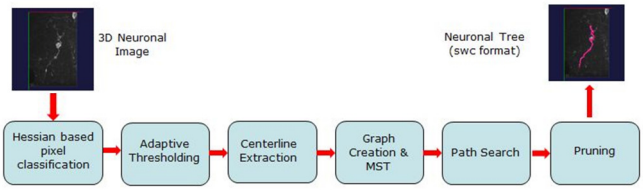


Figure 2: Workflow of Tree2tree-2

The neuron structure is assumed to be tubular, which leads to a model based neuron enhancement technique to delineate the tubular structure from the background. The computation is performed in the scale space paradigm to capture the branches of varying thickness. Inspired by Frangi's method [8], we aim to create a *vesselness* measure to reflect the presence or absence of a neuron. At a voxel $\mathbb{p}$, we have a 3D Hessian (matrix of second partial derivatives) with eigenvalues $|\lambda_1| \leq |\lambda_2| \leq |\lambda_3|$. The tubular portions are characterized by $|\lambda_1| \approx 0, |\lambda_1| \ll |\lambda_2|$ and $\lambda_2 \approx \lambda_3$. The vesselness response at scale σ is obtained as

$$\mathcal{N}_\sigma(\mathbb{p}) = \begin{cases} \dfrac{|\lambda_1(\mathbb{p}) - \lambda_2(\mathbb{p})|^2}{(|\lambda_1(\mathbb{p})||\lambda_2(\mathbb{p}) - \lambda_3(\mathbb{p})|)} & \text{if } \lambda_2(\mathbb{p}), \lambda_3(\mathbb{p}) \leq 0 \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

The hessian enhancement is followed by an adaptive initial segmentation step based on surface evolution to classify the voxels into foreground and background. This variational segmentation scheme poses the problem in an optimization framework that balances the contribution of the signal intensity and the smoothness of the threshold surface. The extracted medial axis of the possibly disjoint binary components represents the sub-parts of the neuronal structure, whose connectivity is established by analyzing the geometric orientation of the components. A minimum spanning tree of the created global graph represents the neuron structure as a graph theoretic tree.

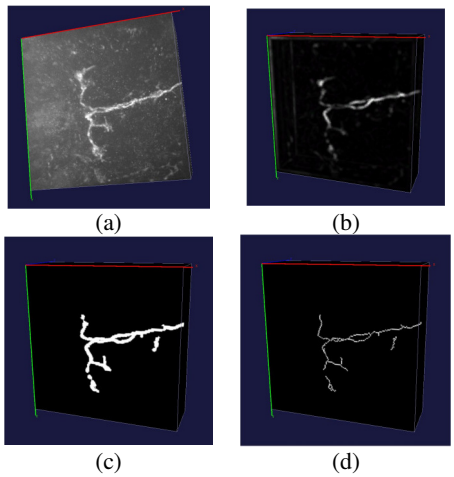


Figure 3: (a) A 3-D confocal stack of a neuron. (b) The vesselness image post hessian enhancement. (c) Segmented neuron and the extracted centerline (d).

The path search method uses a shortest path algorithm to ensure the presence of path between two disjoint components. Once the connectivity information is obtained, the neuron sub-structures are connected by this shortest path, resulting in accurate path connectivity. An automated pruning process is also devised to eliminate any undesired branches, which may appear due to segmentation error. The final neuronal morphology is embedded in 'SWC' file format which contains the geometric and connectivity information of the traced neuron. The SWC file allows visualization of the traced neuron using popular rendering software like Vaa3d. Sample segmentation results are shown in Figure 4. The discussed technique is well equipped to handle complex branching pattern and is completely automated. The algorithm performs well in presence of noise and clutter as shown in Figure 4(c).

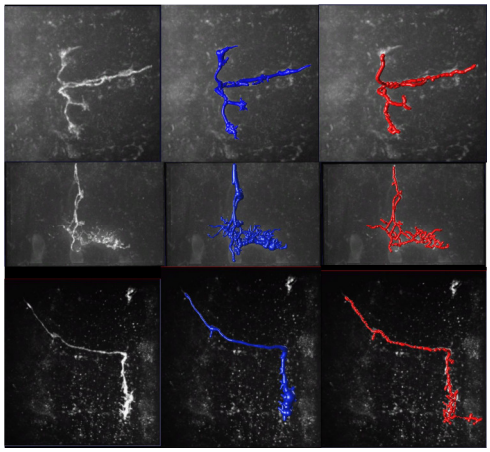


Figure 4: (a) Original confocal stack. (b) Ground truth segmentation using Vaa3d. (c) Segmentation using Tree2Tree.

Apart from the above discussed methods, a number of semi-automated software such as Neuron Studio [9], Simple Neurite Tracer [10] are freely available. A broad class of neuron tracing algorithms is discussed in [11].

3. NEURON MATCHING

Building the neurome requires analysis of thousands of neurons from different species. The first step in achieving this goal is segmentation or tracing. Proper segmentation provides insight into the morphological structure and connectivity of a neuron. The second challenge is that of neuron classification or matching. Given a pre-segmented neuron, an efficient matching algorithm should be able to categorize the neuron into its morphological class. This serves a two-fold purpose. First, this creates a validation mechanism of the tracing algorithm. Second and more importantly, the matching score between two neurons can be used to retrieve a morphologically similar neuron from a vast database. Structural classification is important since it supports Cajal's hypothesis that structure of a neuron is highly correlated with its functionality.

The problem of neuron matching has received relatively less attention over the years. While significant efforts have been made to automate the tracing process, the task of categorizing the neurons still remains an open problem. The DIADEM metric [12] was developed to compute the reconstruction error of a segmented neuron. It uses a branch point proximity based method to calculate the reconstruction error. However, this method is more suitable for comparing multiple reconstructions of the same neuron, since the error metric does not necessarily reflect a substantial difference in structure. The method of Mayerich *et al.* [13] computes the reconstruction error by penalizing the absence of a proper branch or presence of an unwanted loop.

Recently, two approaches were proposed to compute a similarity score between two fully reconstructed neurons. Both the approaches treat a pre-segmented neuron as a graph theoretic tree, with the cell body/soma as its designated root. The first approach *Path2Path* [14] was a novel idea that modeled a neuron as a set of continuous connected paths. Each point in a neuronal path is characterized by a) its *concurrence*, or the number of times the point is shared between the remaining neuronal paths and b) its *hierarchy* that represents the number of levels a point is separated from the root.

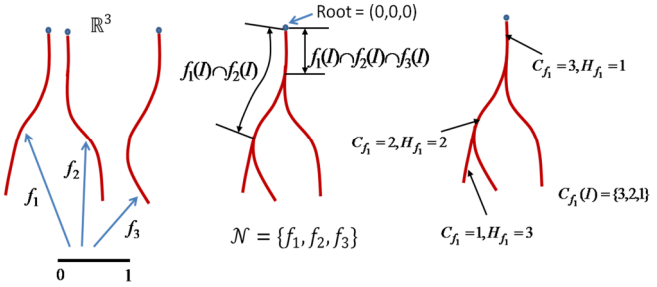


Figure 5: A simulated neuron as a collection of paths.

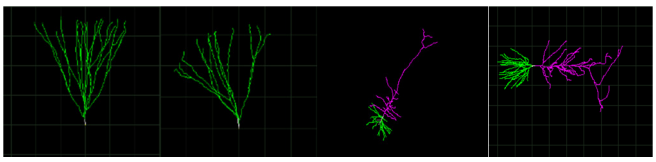
Figure 5 illustrates the concept of concurrence and hierarchy. f_i represents the individual paths. *Concurrence* and *hierarchy* of a path f_i is represented as C_{f_i} and H_{f_i} respectively. The degree of mismatch between the paths f_i and g_j for neurons N and M is given as

$$\varepsilon_{f_i,g_j} = \int_0^1 \frac{|C_{f_i}(t) - C_{g_j}(t)||f_i(t) - g_j(t)|}{\sqrt{H_{f_i}(t)H_{g_j}(t)}}dt \qquad (2)$$

Path2Path is an efficient formulation to compare two neurons. However, to tackle more complicated geometry and branching pattern, another method was introduced to compare neurons based on their local geometric orientation.

The second method highlighted here, the geometric-statistical method [15], computes the shape

histogram of the neuronal points at each level of hierarchy. This formulation reduces the neuron matching problem to computing the distance between two distributions. Also, a penalty term was introduced to account for the mismatch in hierarchies of the neurons. Both *Path2Path* and the geometric-statistical approach were used to compare between neurons of different functionality. While the performance was encouraging for relatively simpler structures, complicated branching patterns create a bottleneck in the classification.



(a) (b) (c) (d)

Figure 6: (a) and (b) are granule cells of the Rat. (c) and (d) shows pyramidal cells of the same species.

For example, in Figure 6, the granule cells (a)-(b) are morphologically similar. However, it is difficult to conclude that the pyramidal cells (c)-(d) are structurally similar, although they serve the same function in the brain. The performance of the present matching algorithms is dependent on the evident structural similarity/dissimilarity between two neurons.

4. FUTURE WORK AND CONCLUSION

From biological perspective, development of the neurome is necessary to analyze the different symptoms related to certain neurological diseases or to measure the amount of degeneration in the neurons with aging. This requires complete automation for the neuron tracing and matching techniques. As mentioned in this paper, both segmentation and matching of neurons is still an open problem. In fact, automated information extraction is the final bottleneck towards achieving the neurome. Despite ongoing research, a large population of researchers in the biological community still depends on manual tools to segment neurons. In fact, most of the successful algorithms in the Diadem challenge (http://diademchallenge.org) were semi-automated. This leaves massive room for research as far as designing automatic tracing algorithm is concerned.

Neuron classification problem is still open. Only initial work has been reported regarding matching and retrieval of neurons. Since the discussed matching methods consider only the path shape and branch orientations of the neurons, it would be interesting to analyze the shape of the neuron contour (2-D) or the neuron surface. The neuron matching can also be performed by devising a sub-graph matching algorithm. However, since subgraph isomorphism is computationally hard, it would be necessary to consider some acceptable approximated algorithms.

REFERENCES

[1] H. Peng, Z. Ruan, F. Long, J. Simpson and E. and G. Myers, "V3D enables real-time 3D visualization and quantitative analysis of large-scale biological image data sets," *Nature Biotechnology,* vol. 28, pp. 348-353, 2010.

[2] H. Peng, F. Long and G. Myers, "Automatic 3D neuron tracing using all-path pruning," *Bioinformatics,* vol. 27, pp. 239-247, 2011.

[3] H. Peng, Z. Z. Ruan, D. Atasoy and S. Sternson, "Automatic reconstruction of 3D neuron structures using a graph-augmented deformable mode," *Bioinformatics,* vol. 26, pp. 38-46, 2010.

[4] Y. Wang, A. Narayanaswamy, C.-L. Tsai and B. Roysam, "A Broadly Applicable 3-D Neuron Tracing Method Based on Open-Curve Snake," *Neuroinform,* pp. 193-217, 2011.

[5] C. Xu and J. L. Prince, "Snakes, Shapes, and Gradient Vector Flow," *IEEE Transactions on Image Processing,* pp. 359-369, 1998.

[6] S. Basu, A. Aksel, B. Condron and S. T. Acton, "Segmentation and tracing of neurons in 3D," *IEEE Transactions on Info. Tech. Biomedicine,* in press, 2012.

[7] S. Mukherjee, S. Basu, B. Condron and S. T. Acton, "Tree2Tree2: Neuron tracing in 3D," in *IEEE ISBI(accepted)*, 2013.

[8] A. Frangi, W. Niessen, K. Vincken and M. Viergever, "Multiscale vessel enhancement filtering," *MICCAI,* vol. 1496, pp. 130-137, 1998.

[9] S. Wearne and a. et, "New Techniques for imaging, digitization and analysis of three-dimensional neural morphology on multiple scales," *Neuroscience,* pp. 661-680, 2005.

[10] M. H. Longair, D. A. Baker and J. D. Armstrong, "Simple Neurite Tracer: Open Source software for reconstruction, visualization and analysis of neuronal processes," *Bioinformatics,* vol. 27, no. 17, pp. 2453-2454, 2011.

[11] E. Meijering, "Neuron Tracing in Perspective," *Cytometry Part A,* vol. 77, pp. 693-704, 2010.

[12] T. Gillette, K. Brown and G. Ascoli, "The DIADEM metric: comparing multiple reconstructions of the same neuron.," *Neuroinformatics,* pp. 233-245, 2011.

[13] D. Mayerich, C. Björnsson, J. Taylor and B. Roysam, "Metrics for comparing explicit representations of interconnected biological networks," *BioVis,* pp. 79-86, 2011.

[14] S. Basu, B. Condron and S. T. Acton, "Path2Path: Hierarchical path-based analysis for neuron matching," in *International Symposium on Biomedical Imaging*, Chicago, 2011.

[15] S. Mukherjee, S. Basu, B. Condron and S. Acton, "A geometric-statistical approach toward neuron matching," in *International Symposium on Biomedical Imaging*, Barcelona, 2012.