VESSELSHIFT: A MEAN-SHIFT BASED METHOD FOR NEURITE TRACING

Arturo Deza, Aruna Jammalamadaka, B.S. Manjunath

Department of Electrical and Computer Engineering, University of California, Santa Barbara
arturodeza@gmail.com, arunaj@ece.ucsb.edu, manj@ece.ucsb.edu

ABSTRACT

A novel approach for detection of neurite trace nodes from 3D confocal microscopy images is presented. Many neurite tracing methods rely heavily on an initial seed point or trace node detection step upon which the neurite trace is built. This step is crucial, as improper node detection results in a loss of information that cannot be re-acquired with commonly used subsequent tracing steps such as minimum spanning trees and active contours. We propose an improved trace node detection method dubbed "Vesselshift" which first identifies possible nodes along the centerline of the vessel through 3D Gradient Vector Flow and then mean-shift clusters these points to produce accurate trace nodes. The method is less dependent on accurate segmentation of the neurite than previous methods, achieving robustness in uncertain regions common with this type of data. We present qualitative and quantitative results on the DIADEM Olfactory Projection Fiber dataset.

Index Terms— neuronal tracing, trace node detection, mean-shift, gradient vector field, DIADEM dataset

1. INTRODUCTION

Recent advances in imaging technologies and affordable computational power have made the acquisition of 3D neuronal data more accessible, and as the amount of this data increases so has the necessity for automated 3D algorithms for tracing and modeling neuronal structures. However, the attempts of automated algorithms to satisfy scientists' desires to understand correlation between dendrite physiology and morphology have not yet reached a state of usability [1, 2]. Thus the DIADEM Challenge [3] was created as a benchmark for neuronal reconstruction algorithms, in order to further the state of the art in a well-established manner. The challenge presents 6 diverse datasets as well as a scoring metric for evaluating algorithm performance quantitatively. The finalist algorithms from this challenge along with most other neurite tracing algorithms, share a few standard processing steps. These steps include segmentation of the neurite structure, creation of nodes along the segmented neurite centerline, imposition of geometrical constraints such as trace smoothness and width consistency, and finally edge linking in such a way that the trace constructs a binary tree which can be output in the standard SWC format (details on SWC format can be found in [3]). Of these steps one of the most important and sensitive steps is that of node detection, as nodes which are missed in this step cannot be re-acquired in later steps. Ideal trace nodes lie along the neurite centerline and span all branches such that when connected a trace of the neurite is formed.

In this paper we present a novel technique for finding such nodes as an essential step in the tracing pipeline. Our contributions include the method's robustness towards faulty segmentations as well as it's generality, allowing it's usage as a substitute node detection step which could potentially improve the quality of any existing neurite tracing algorithms. The paper is organized as follows: Section 2 gives a brief overview of existing state-of-the-art algorithms and the importance of node quality in each technique, Section 3 outlines the Vesselshift framework, and finally Section 4 concludes with both qualitative and quantitative comparisons to existing methods.

2. PREVIOUS WORK

Most tracing algorithms [4, 5, 6, 7, 8], including DIADEM finalist algorithms [9, 10, 11, 12, 13], contain variations of these general steps: segmentation, node detection, and connection of those nodes according to semantic geometrical constraints. Segmentation can range from simple binarization [11], to cylindrical filtering with thresholds [12, 10], to 2D graph-cut algorithms [9]. Node detection algorithms generally involve centerline extraction based on that segmentation [4, 13, 11, 9] or selection of highest intensity filtered points [10, 12]. Connection of nodes is usually performed by creating a constrained Minimum Spanning Tree (MST) [10, 11, 12, 13]. Nodes can also be used as seed points for active contours [9].

Of the algorithms mentioned, many [7, 9, 11, 8] are segmentation dependent. This is a problem with bioimages because image quality is commonly low, due to poor imaging conditions or staining. Instead of relying on the segmentation for centerline extraction, or moving raw nodes from a filtered image along the Gradient Vector Flow (GVF) [14] our method uses the GVF divergence of the $original$ image data to directly acquire an initial set of centerline nodes.

This work was funded by ProUNI-2547 & NSF award III-0808772

Preprint submitted to 2013 International Symposium on Biomedical Imaging: From Nano to Macro. Received November 7, 2012.

In addition, rather than sampling GVF points for placement of nodes along the centerline, our algorithm uses mean shift to find node cluster centers as a smarter way of sampling the initial GVF points. The nodes are then loosely validated by the segmentation to avoid extreme outlier false positive nodes induced by noise. We choose to compare our method to the methods described in [9] and [7] due to their readily available state-of-the-art results on the Olfactory Projection Fiber Dataset. It should be noted however that the tracing method outlined can be applied to any long vessel-like bio-structures, such as any of the other 5 DIADEM datasets as well as blood vessels, with very little parameter tuning.

3. PROPOSED METHOD

3.1. Background

3.1.1. Node Detection

The 3-dimensional Gradient Vector Flow field [14] is defined as the vector field $\mathbf{v}(x,y,z) = (u(x,y,z), v(x,y,z), w(x,y,z))$ that minimizes the energy functional E in equation (1), where f is the image and μ is the weighting parameter.

$$E = \iiint \mu(u_x^2+u_y^2+u_w^2+v_x^2+v_y^2+v_z^2+w_x^2+w_y^2+w_z^2) + |\nabla f|^2|\mathbf{v}-\nabla f|^2 dxdydz \quad (1)$$

Once $\mathbf{v}$ has been computed, we define the divergence function $Q \in \mathbb{R}^d$ as $Q = \nabla \cdot \mathbf{v}$. From this we get the local minima or "sink points" $\mathbf{s} = \{\mathbf{s_1}, ..., \mathbf{s_N}\} \in S = \mathbb{R}^d$ that satisfy $Q(\mathbf{s_i}) \leq \nu$ where ν is a threshold parameter. Notice that ν has a negative value, as a negative divergence indicates that the vector field is converging. N, the number of sink points selected, decreases as $|\nu|$ increases.

3.1.2. Mean-shift Clustering

In our framework we use the $Uniform\ Kernel$ $\mathbf{G}_U$ as our mean-shift kernel to compute the mean-shift $\mathbf{M}_{\lambda,\mathbf{G}}$, and $g(x)$ as the profile function of $\mathbf{G}_U$ [16] as shown in equation (2):

$$\mathbf{M}_{\lambda,\mathbf{G}} = \frac{\sum_{i=1}^{N} g\left(\left|\left|\frac{\mathbf{x}-\mathbf{s_i}}{\lambda}\right|\right|^2\right)\mathbf{s_i}}{\sum_{i=1}^{N} g\left(\left|\left|\frac{\mathbf{x}-\mathbf{s_i}}{\lambda}\right|\right|^2\right)} - \mathbf{x} \quad (2)$$

where λ is the bandwidth parameter, $\mathbf{x}$ is the kernel's position, and $\mathbf{s_i}$ are the $1 \rightarrow N$ sink points previously defined.

3.2. Vesselshift Algorithm

The Vesselshift pipeline begins by computing the 3D GVF of the image volume. It then calculates the divergence of the GVF and the matrix is thresholded by ν. Fig.1. (b) shows a 2D projection of the resulting sink points around the centerline. Next a simple marker-based watershed segmentation is used to eliminate points due to noise in the image. Markers are obtained from using a Frangi filter [17] and taking highest intensity points. Finally the points are windowed and mean-shift clustered, setting the bandwidth parameter λ (2) as a pseudo-average sampling distance.

Algorithm 1 Vesselshift(λ, ν)

Load Original Image Volume
Compute GVF of the Original Image Volume
Get first set of sink points $\{\mathbf{s}\}$ from the image f:
for every voxel $\mathbf{f_i} \in f$ **do**
  if $Q(\mathbf{f_i}) \leq \nu$ **then**
    $\{\mathbf{s}\} \leftarrow \{\mathbf{s}\} + \mathbf{f_i}$
  end if
end for
Get Marker-based watershed Segmentation X
Validate sink points with Segmentation X:
for $\mathbf{s_i} \in \{\mathbf{s}\}$ **do**
  if $\mathbf{s_i} \in X$ **then**
    $\{\mathbf{s^x}\} \leftarrow \{\mathbf{s^x}\} + \mathbf{s_i}$
  end if
end for
Apply mean-shift to $\{\mathbf{s^x}\}$ to get final nodes $\{\mathbf{q}\}$:
$\{\mathbf{q}\} = mean\text{-}shift_\lambda(\{\mathbf{s^x}\})$

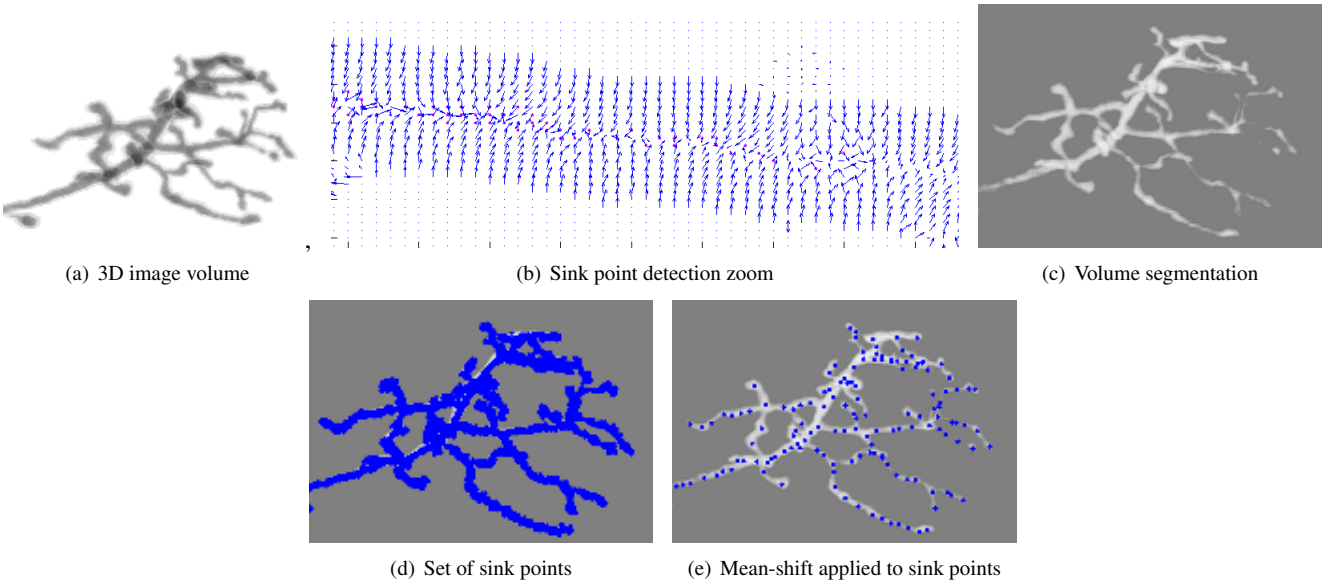
3.3. Parameter Selection

For tracing it is important that the mean-shift cluster centers be close to the centerline, therefore we use bandwidth $\lambda = k_0\mathbf{D}$ where $\mathbf{D}$ is the average diameter of the structure to be segmented. If k_0 is too small, the final nodes $\{\mathbf{q}\}$ will contain unwanted spurs. Conversely if k_0 is too big, then critical nodes, such as branching points or those in thin branches might be discarded. In our experiments we obtain a maximum DIADEM metric score [2] using $k_0 = 0.9$, GVF weighting parameter $\mu = 0.15$ and divergence threshold $\nu = -(2/3)(max(|Q|))$ which in our case came out to -2.0. These 3 parameters can easily be tuned for other datasets if needed.

4. RESULTS AND DISCUSSION

To validate our algorithm we have compared our nodes with the nodes output from Lee et al.'s skeletonization-based tracing algorithm [7]. In Table 1, we show that Vesselshift outperforms Lee et al.'s method in every Olfactory Projection Fiber dataset that they have used. The $Dis(N_2, N_1)$ score they use indicates the average minimum distance from each node in reconstruction N_2 to the closest node in the Gold Standard reconstruction N_1, thus a perfect $Dis(N_2, N_1)$ score would be 0. All scores are presented with a $\pm\ \sigma$ error.

Preprint submitted to 2013 International Symposium on Biomedical Imaging: From Nano to Macro. Received November 7, 2012.



(a) 3D image volume (b) Sink point detection zoom (c) Volume segmentation

(d) Set of sink points (e) Mean-shift applied to sink points

Fig. 1. The Vesselshift pipeline illustrated by the figures above. (a) Original 3D image volume. (b) Divergence of GVF. Sink points shown in magenta. (c) Watershed segmentation. (d) Sink points shown on image volume. (e) Final Vesselshift nodes. Images were rendered in BioView3D [15]

Table 1. Comparative $Dis(N_2, N_1)$ scores

Data	Size	Lee et al.	Vesselshift
OP_1	512 x 512 x 60	1.41	**1.40 ± 0.01**
OP_6	512 x 512 x 101	1.87	**1.84 ± 0.01**
OP_7	512 x 512 x 71	1.75	**1.72 ± 0.01**
OP_9	512 x 512 x 92	1.65	**1.31 ± 0.01**

Table 2 shows a comparison using DIADEM metric scores [2] with Wang et al.'s active contour method [9]. Because their algorithm initially detects seed points for active contours instead of trace nodes we must compare their output trace to Vesselshift nodes connected by a simple MST. The MST we use lacks geometrical or semantic constraints and is simply created by inputting the given root node per dataset, the detected Vesselshift nodes, and an edge weighting function given by

$$w_V(i,j) = \sum_{k=1}^{m}((1 - X_{(x_k,y_k,z_k)}) + 1) \qquad (3)$$

where m is the integer number of voxels in the euclidean distance between nodes $\{\mathbf{q_i}$ and $\mathbf{q_j}\}$, and $X_{(x_k,y_k,z_k)} = 1$ if a voxel is considered to be on the neurite, or in the foreground of the segmentation. As shown in the table, even this very simplistic tracing algorithm outperforms Wang et al. on all but 2 of the OP datasets.

Fig. 2 shows a qualitative comparison between the algorithms in order to further analyse the reasons for this improvement. The node spacing in [9] is more uniform due to their sampling method, but their trace fails to include two important branches marked by the two red-shaded rectangles in Fig 2(b). This is because the branches shown in green were disconnected from the main arbor in the segmentation step and therefore the corresponding active contours could not later be merged. Although Vesselshift node spacing is not as uniform in comparison, we are able to avoid the pre-segmentation step which delimits the active contour expansion boundaries and has the potential to propagate unrecoverable errors to the later steps of the tracing algorithm.

Table 2. Comparison of DIADEM scores

Data	Size	Wang et al.	Vesselshift
OP_1	512 x 512 x 60	0.860	**0.860 ± 0.08**
OP_2	512 x 512 x 88	0.853	**0.858 ± 0.00**
OP_3	512 x 512 x 62	0.652	**0.725 ± 0.03**
OP_4	512 x 512 x 67	0.790	0.602 ± 0.19
OP_5	512 x 512 x 76	0.316	**0.865 ± 0.01**
OP_6	512 x 512 x 101	0.772	0.687 ± 0.00
OP_7	512 x 512 x 71	0.826	**0.836 ± 0.03**
OP_8	512 x 512 x 85	0.792	**0.834 ± 0.05**
OP_9	512 x 512 x 92	0.680	**0.708 ± 0.02**

We conclude that Vesselshift is an ideal candidate for node detection in the neuron tracing pipeline. Additionally, it could also be used in tracing algorithms that require an initial seed detection phase [9] at the pre-processing phase, for later usage in a non-MST based workflow. Intended future work would be to complement Vesselshift with geometrical constraints for the MST weighting function, to design a more robust end-to-

Preprint submitted to 2013 International Symposium on Biomedical Imaging: From Nano to Macro. Received November 7, 2012.

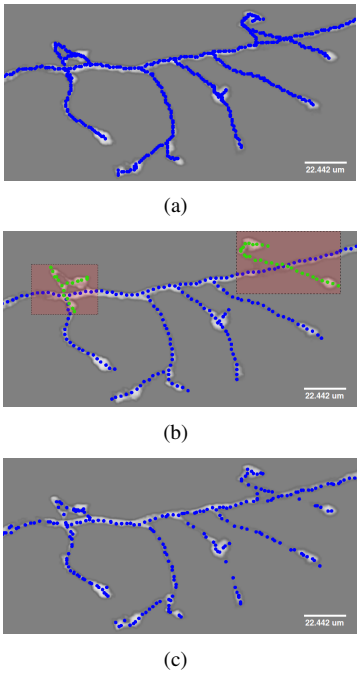


Fig. 2. (a) Gold Standard Reconstruction. (b) Wang et al.'s output trace (c) Vesselshift nodes.

end automated tracing model.

5. REFERENCES

[1] M. Halavi, K.A. Hamilton, R. Parekh, and G.A. Ascoli, "Digital reconstructions of neuronal morphology: three decades of research trends," *Frontiers in Neuroscience*, vol. 6, 2012.

[2] T.A. Gillette, K.M. Brown, and G.A. Ascoli, "The diadem metric: comparing multiple reconstructions of the same neuron," *Neuroinformatics*, vol. 9, no. 2, pp. 233–245, 2011.

[3] Howard Hughes Medical Institute, "DIADEM Challenge," `www.diademchallenge.org`, 2012, [Online; accessed August 10th, 2012].

[4] AR Cohen, B. Roysam, and JN Turner, "Automated tracing and volume measurements of neurons from 3-d confocal fluorescence microscopy data," *Journal of microscopy*, vol. 173, no. 2, pp. 103–114, 2011.

[5] H. Peng, Z. Ruan, D. Atasoy, and S. Sternson, "Automatic reconstruction of 3d neuron structures using a graph-augmented deformable model," *Bioinformatics*, vol. 26, no. 12, pp. i38–i46, 2010.

[6] K.A. Al-Kofahi, S. Lasek, D.H. Szarowski, C.J. Pace, G. Nagy, J.N. Turner, and B. Roysam, "Rapid automated three-dimensional tracing of neurons from confocal image stacks," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 6, no. 2, pp. 171–187, 2002.

[7] P.C. Lee, C.C. Chuang, A.S. Chiang, and Y.T. Ching, "High-throughput computer method for 3d neuronal structure reconstruction from the image stack of the drosophila brain and its applications," *PLOS Computational Biology*, vol. 8, no. 9, pp. e1002658, 2012.

[8] M. Sofka and C.V. Stewart, "Retinal vessel centerline extraction using multiscale matched filters, confidence and edge measures," *Medical Imaging, IEEE Transactions on*, vol. 25, no. 12, pp. 1531–1546, 2006.

[9] Y. Wang, A. Narayanaswamy, C.L. Tsai, and B. Roysam, "A broadly applicable 3-d neuron tracing method based on open-curve snake," *Neuroinformatics*, vol. 9, no. 2, pp. 193–217, 2011.

[10] E. Türetken, G. González, C. Blum, and P. Fua, "Automated reconstruction of dendritic and axonal trees by global optimization with geometric priors," *Neuroinformatics*, vol. 9, no. 2, pp. 279–302, 2011.

[11] P. Chothani, V. Mehta, and A. Stepanyants, "Automated tracing of neurites from light microscopy stacks of images," *Neuroinformatics*, vol. 9, no. 2, pp. 263–278, 2011.

[12] T. Zhao, J. Xie, F. Amat, N. Clack, P. Ahammad, H. Peng, F. Long, and E. Myers, "Automated reconstruction of neuronal morphology based on local geometrical and global structural models," *Neuroinformatics*, vol. 9, no. 2, pp. 247–261, 2011.

[13] E. Bas and D. Erdogmus, "Principal curves as skeletons of tubular objects," *Neuroinformatics*, vol. 9, no. 2, pp. 181–191, 2011.

[14] C. Xu and J.L. Prince, "Gradient vector flow: A new external force for snakes," in *Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on*. IEEE, 1997, pp. 66–71.

[15] K. Kvilekval, D. Fedorov, B. Obara, A. Singh, and BS Manjunath, "Bisque: a platform for bioimage analysis and management," *Bioinformatics*, vol. 26, no. 4, pp. 544–552, 2010.

[16] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, no. 5, pp. 603–619, 2002.

[17] A. Frangi, W. Niessen, K. Vincken, and M. Viergever, "Multiscale vessel enhancement filtering," *Medical Image Computing and Computer-Assisted Interventation-MICCAI98*, pp. 130–137, 1998.

Preprint submitted to 2013 International Symposium on Biomedical Imaging: From Nano to Macro. Received November 7, 2012.