Multiple Structure Tracing in 3D Electron Micrographs

Vignesh Jagadeesh, Nhat Vu, and B.S.Manjunath *

Department of ECE and Center for Bioimage Informatics
University of California, Santa Barbara CA - 93106

Abstract. Automatic interpretation of Transmission Electron Micrograph (TEM) volumes is central to advancing current understanding of neural circuitry. In the context of TEM image analysis, tracing 3D neuronal structures is a significant problem. This work proposes a new model using the conditional random field (CRF) framework with higher order potentials for tracing multiple neuronal structures in 3D. The model consists of two key features. First, the higher order CRF cost is designed to enforce label smoothness in 3D and capture rich textures inherent in the data. Second, a technique based on semi-supervised edge learning is used to propagate high confidence structural edges during the tracing process. In contrast to predominantly edge based methods in the TEM tracing literature, this work simultaneously combines regional texture and learnt edge features into a single framework. Experimental results show that the proposed method outperforms more traditional models in tracing neuronal structures from TEM stacks.

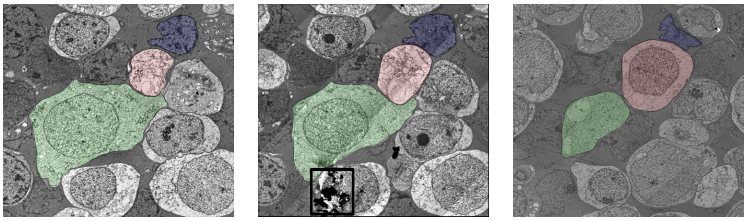
Keywords: Tracing, CRFs, Textures

1 Introduction

Understanding the interconnectivity structure of the brain is a grand challenge in neuroscience. Recent developments in imaging have enabled capturing massive amounts (in terabytes) of Transmission Electron Micrograph (TEM) data at sub-nanometer resolutions. Manual analysis of these data repositories is infeasible, justifying the need for evolving image analysis algorithms. One of the challenges in interpreting these data repositories lies in automated tracing of multiple interacting neuronal structures in 3D. This paper proposes a robust and efficient tracing algorithm employing conditional random fields(CRFs).
Problem Definition: Images displayed in Figure 1 correspond to different z-slices from a TEM stack. Three sample structures are shaded to illustrate difficulty levels in tracing: simple(red), medium(green) and hard(blue). As can be seen, the structures express regional textures that are discriminative and noisy.

* This work was supported by NSF OIA 0941717. The authors thank Dr.Robert Marc, Dr.Brain Jones and Dr.James Anderson from the Univ. of Utah for providing data used in experiments and for useful discussions.

Furthermore, each structure has special gradient profiles that cannot be captured using simple gradient operators. Finally, structures could deform considerably without arbitrary movement from one slice to another. The figure also illustrates problems introduced during the imaging process (sample damage shown by a black box in Figure 1(b), illumination artifacts from Figure 1(b) to 1(c)). As is evident from Figure 1, it would be desirable to model regional cues (in-



(a) (b) (c)

Fig. 1. (Best Viewed in Color) Three representative slices with manual annotations. Structure in red is simple to trace since it has well defined gradients and regional cues. Structure in green is tougher since it shares boundaries with many different neighbors and a variety of gradient profiles arise on its border. Structure in blue is toughest to trace due to inconsistent region and edge information.

tensity, texture etc ..), characteristic gradient profiles (learnt edges) and 3D smoothness constraints (dynamics of structures in the z-direction) in a unified framework that implicitly handles arbitrary changes in topology. This work aims to convince the reader that CRFs are capable of exploiting multiple information sources for tracing neuronal structures in TEM stacks.

Contributions: The primary contribution of this work is in designing an algorithm for tracing multiple interacting 3D structures in TEM stacks. Salient features of the proposed model include:

- Semi-supervised edge learning scheme for propagating high confidence edge maps during the tracing process (Section 2.1)
- Design of a higher order CRF cost for simultaneously enforcing 3D label smoothness and capturing regional textures (Section 2.2)

Related Work: While there has been a lot of interest in EM tracing over the past year, this work differs from existing literature [3],[6],[5],[4] (and references therein) in the following ways. The nature of data considered is vastly different, in that it has discriminative regional texture in addition to characteristic gradient profiles. Most existing techniques approach the problem employing strong edge learning methods since the datasets considered present edge detection as the important challenge. In particular [7] propose a global technique that performs optimization over the entire stack. However, scaling global techniques to large datasets would be difficult. Sample damage and unfavorable imaging conditions could further seriously affect a truly global approach. This motivates the

proposed approach where multiple interacting structures are traced in a scalable manner.

The proposed solution adopts a CRF framework for integrating multiple sources of information described previously. The following paragraph provides a brief background and introduces notations that will be used in this paper. CRFs are image models that capture contextual interaction between pixels. Given features x_{p^z} of a pixel p in slice z of the stack, the goal is to infer its label $y_{p^z} \in \mathcal{L} = \{1, 2..L\}$. $\mathcal{L}$ is the label set containing L labels, which in the present case is the number of structures to be traced. $\mathcal{P}^z$ denotes the set of pixels constituting slice z. Labeling is achieved by minimizing a CRF energy comprising unary, interaction and higher order terms, see Equation 1. Unary potentials encode the likelihood of a pixel p^z to take on label y_{p^z}. Interaction potentials encourage label smoothness between p^z and its neighbors q^z contained in a neighborhood system $\mathcal{N}_{p^z}$. Higher order terms encourage label consistency across any clique (group of pixels) c^z contained in the set of cliques $\mathcal{C}^z$ in slice z. For this work, cliques are superpixels generated by oversegmentation. The CRF energy is given by:

$$E(y_{p^z}) = \underbrace{\sum_{p^z \in \mathcal{P}^z} V_p(y_{p^z})}_{Unary\ Potential} + \underbrace{\sum_{p^z \in \mathcal{P}^z, q^z \in \mathcal{N}_{p^z}} V_{pq}(y_{p^z}, y_{q^z})}_{Interaction\ Potential} + \underbrace{\sum_{c^z \in \mathcal{C}^z} V_c(y_{c^z})}_{Higher\ Order\ Term} \quad (1)$$

The above energy function can be efficiently minimized using graph cuts if each term obeys a submodularity constraint. The time complexity of inference employed in this work is similar to traditional alpha expansions. For further details [9],[2],[1] are comprehensive sources of reference.

This paper is organized as follows. The second section describes the construction of potential functions for the proposed higher order CRF model. The subsequent section presents experimental results on TEM stacks with a quantitative analysis. The final section concludes the paper with a discussion on future work.

2 Multiple Structure Tracing with Higher Order CRFs

The proposed model is obtained by constructing unary, interaction and higher order terms in Equation 1. The unary potentials are coarse object/background likelihoods computed using information from previous segmentations. The unary potentials are unregularized and require interaction models (second and higher order) for smooth segmentation. Traditional second order interaction models comprise first order gradient operators, which are often incapable of capturing a wide range of edge profiles. In contrast, the proposed semi-supervised scheme for edge propagation captures a wider range of edge profiles using a learnt model as shown in Figure 2. The edge propagation scheme does not assume smoothness across the z-direction in its construction. However, there are scenarios where unary and pairwise terms become unreliable (for the hard structure in Figure 1(c) and in case of sample damage). In such scenarios, a model for capturing higher

order regional interactions for resisting failure caused due to first and second order models, and for enforcing smoothness across the third dimension is required. The robust Pn model is employed for this purpose.

2.1 Semi-Supervised Edge Propagation

Interaction potentials are given by $V_{pq}(y_{p^z}, y_{q^z}) = \lambda_I \exp(-\frac{(I_{p^z} - I_{q^z})^2}{2\sigma_I^2}) \frac{1}{dist(p^z, q^z)}$ $\delta(y_{p^z} = y_{q^z})$. λ_I, σ_I are parameters controlling influence of interaction potentials and edge quality respectively, δ is a dirac delta function evaluating to 0 if the condition in parenthesis is satisfied and 1 otherwise. Computing V_{pq} in the above fashion is equivalent to a first order gradient operator, which is not suitable for TEM datasets. This work leverages contour initialization as partial labeling and propagates edges through the stack using a semi supervised scheme. The idea is to learn edge textons [10] to improve structure specific edge detection. This is in contrast to recent efforts that learn application specific edges using large amounts of training data. For ease of explanation, a two label problem with the goal of propagating edges between slices $z-1$ and z is considered.
The filter bank considered comprises of M filters and is given by $\mathcal{F} = [F_1 F_2 ... F_M]$,

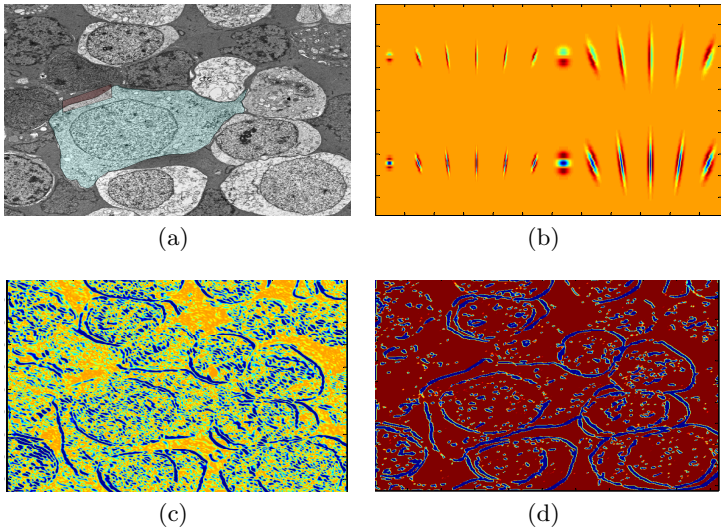


Fig. 2. (Best viewed in color, red pixels have high values/blue pixels have lower values) (a) Green region is to be traced. Red region in (a) shows a sample edge that needs to be detected. (b) Filterbank used (c) Traditional second order energy, gradients not well localized and edges in the red region shown in (a) are missed. (d) Proposed scheme provides accuracy localization for structure of interest. Observe red region in (a) is detected with high confidence.

see 2(b). Specifically, the filter bank has twelve first and twelve second deriva-

tive filters, each at six orientations and two scales (a subset of the Leung Malik filter bank). The labels $y_{p^{z-1}}$ are known and so are edge pixels for slice $z-1$. Initially the images are convolved with $\mathcal{F}$ resulting in filter responses $R^i_{p^{z-1}}$ and $R^i_{p^z}$ for slices $z-1$ and z respectively. The intuition behind using a filter bank is capture varied gradient profiles exhibited by the individual filters, in contrast to simple first order gradients. Since there are M filter responses, concatenation of all responses $R^i_{p^{z-1}}, \forall i \in [1,2,..M]$ at a pixel p^{z-1} is represented by the M dimensional vector $\mathbf{R}_{p^{z-1}}$. Subsequently, the edge filter responses are clustered in an M dimensional feature space for the the K edge textons $\mathbf{C}_k$. The texton indices t_{p^z} are found by performing a nearest neighbor search of filter responses on the image in slice z with the textons $\mathbf{C}_k$. The feature descriptor for training classifiers are local histograms of texton indices, $h_{p^z}(k) = \sum_{q^z \in \mathcal{N}_{p^z}} \delta(t_{q^z} \neq k), \forall k \in [1,2,...K], h_{p^z} \in \mathcal{R}^K$. A naive Bayes classifier is trained using the feature vectors $h_{p^{z-1}}$ and tested on features from the current frame h_{p^z}. The output of the classifier is a posterior probability $Pr(e_{p^z})$ that yields large values at locations where there is a high probability of the structure specific edge. The interaction potentials for the conditional random field can now be rewritten as $V_{pq}(y_{p^z}, y_{q^z}|y^{z-1}) = \lambda_I(1 - \frac{Pr(e_{p^z})+Pr(e_{q^z})}{2})\delta(y_{p^z} = y_{q^z})$.

2.2 Robust Higher Order Potentials

It is well known that higher order CRFs are capable of modeling larger spatial interactions. These models are ideal for capturing textures inherent in the neuronal structures. This work adopts the recently proposed Robust P^n model[8], a class of potential functions that are a strict generalization of the Potts model for second order interactions. The idea behind the model is that pixels constituting a superpixel (homogenous regions) are more likely to take the same label. The cost is expressed as $V_c(y_{c^z}) = \min\{\min_{k \in \mathcal{L}}((|c^z| - n_k(y_{c^z}))\theta_k + \gamma_k), \gamma_{max}\}$, where $|c^z|$ is the cardinality of the clique in slice z, $\mathcal{L}$ is the label set containing the set of possible labels, $n_k(y_{c^z})$ is the number of labels in the clique taking label $k \in \mathcal{L}$, $\theta_k = \frac{\gamma_{max}-\gamma_k}{Q}$ and Q is a truncation parameter controlling magnitude of label violations in the clique. γ_k, γ_{max} are penalties associated with the clique taking label k and mixed labeling respectively. γ_k is usually set to zero since uniform superpixel labelings are not penalized, $\gamma_{max} = |c|^{\theta_\alpha}(\theta^h_p + \theta^h_v G(c))$, where $\theta_\alpha, \theta^h_p, \theta^h_v$ are free parameters and $G(c^z)$ is a quantity indicative of superpixel quality. While enforcing label consistency, it is imperative that smoothness across the z-direction is preserved. One requires a cost $V_c(y_{c^z}|y^{z-1})$ and a method for evaluating G_c for superpixels.

Intuition: If a contour with label l propagates down till slice $z-1$, then the superpixels in slice z overlapping with the contour in slice $z-1$ are most likely to take the label l. In other words, the dominant label for superpixel c^z (overlapping with contour labelled l in slice $z-1$)is l. If all pixels constituting the superpixel take on the dominant label, minimal penalty is incurred. On the other hand, as the number of pixels violating the dominant label increases, higher cost (closer to γ_{max}) is incurred.

Modified Higher Order Cost: The segmentation of the current slice needs to respect the label homogeneity of the current slice, and also preserve smoothness across the z-direction. The proposed model incorporates smoothness across the z-direction using the variables γ_k. In previous constructions, the value of γ_k was usually set to zero since label homogeneity indicated smooth segments in 2D. However, label homogeneity across 2D does not directly imply smoothness across 3D. In particular, an unsymmetrical Dice coefficient is used for measuring the overlap of a certain superpixel (c^z) to its overlapping contour in slice $z-1$.

$$n^k_{c^z} = \frac{\sum_{p^z \in P^z} M^{c^z}_{p^z} \wedge \delta(\hat{y}_{p^z} \neq k)}{\sum_{p^z \in P^z} M^{c^z}_{p^z}}, \quad M^{c^z}_{p^z} = \begin{cases} 1, & \forall p^z \in c^z, \\ 0, & otherwise \end{cases}$$

In the above equation $\hat{y}_{p^z} = y_{p^{z-1}}$, meaning that label predictions for slice z are labels propagated from slice $z-1$. Smoothness across the third dimension can now be incorporated as $\gamma_k \propto 1 - \frac{n^k_{c^z}}{|c^z|}, \quad c^z \in \mathcal{C}^z, \quad \forall k \in \mathcal{L}$. The constraint $\gamma_k < \gamma_{max}$ is always enforced for the cost to be minimized by graph cuts. Superpixel quality is evaluated using the variance of local intensity features on the superpixel. The modified higher order cost is obtained by substituting for γ_k, $V_c(y_{c^z}|y^{z-1}) = \min\{\min_{k \in \mathcal{L}}((|c^z| - n_k(y_{c^z}))\theta_k + \gamma_k), \gamma_{max}\}$.

2.3 First Order Potentials

The unary potential models the likelihood of a certain pixel taking up label $l \in \mathcal{L} = \{1, 2, ..L\}$. Electron Micrograph data is rich in texture, but not of the sort one would find in traditional texture analysis literature. It is used as a valuable cue by biologists, but seems to contain a lot of noise. A local multiscale feature similar to texture histograms, $I^i_z = I_z * g_{\sigma_i}, \quad 1 \leq i \leq N_f$ where $N_f = 3$ is employed. The above equation refers to smoothing of image I_z at position z on the stack by a Gaussian kernel g_{σ_i} with variances σ_i. Concatenation of filter responses at each pixel yields a feature vector in $\mathcal{R}^{N_f}$ and likelihoods are obtained by standard histogram backprojection techniques. The overall unary potential can be expressed as: $V_p(y_{p^z}|y_{p^{z-1}}, I_z) = -\log(Pr(I_z|y_{p^z}, y_{p^{z-1}})Pr(y_{p^{z-1}}|y_{p^z}))$. Note that $Pr(I_z|y_{p^z}, y_{p^{z-1}})$ is obtained by backprojecting the multiscale histograms, and $Pr(y_{p^{z-1}}|y_{p^z})$ is a signed distance function constraining the contour to be close to its position in the previous slice (shape prior). Further, optical flow fields could also be employed if there are mild registration errors between slices for simultaneous segmentation and registration.

3 Experimental Results

Experimental results are reported on Electron Micrographs of the retina. Neuronal structures were traced for over 45 slices of the stack. Results are compared with ground truth for a quantitative analysis of pixel errors.

Single Structure Tracing: Tracing single structures is performed to provide

a proof of concept for the edge learning and higher order models. Tracing was done by providing an initial contour on the first frame of the stack, and is quantified by computing the F-measure (where P and R are precision and recall), $F = \frac{2PR}{P+R}$. As can be observed from Figure 3(b), traditional second order terms get distracted by noisy gradients and were not able to recover once the contour was lost. The error on subsequent slices were additive, leading to poor performance as shown by the red lines. However, the second order model with proposed edge learning was able to resist distractions of noisy gradients as can be seen from green lines. On the medium difficulty structure, average F-measure for the traditional scheme was 0.925, in comparison to 0.962 for the proposed scheme. Figure 3(c) illustrates an example using the difficult target where the higher order model was able to outperform simple second order interactions. On the hard structure, average F-measure for the second order model was 0.817, while the higher order model yielded 0.889.

Multiple Structure Tracing: Experiments were also performed for multi-

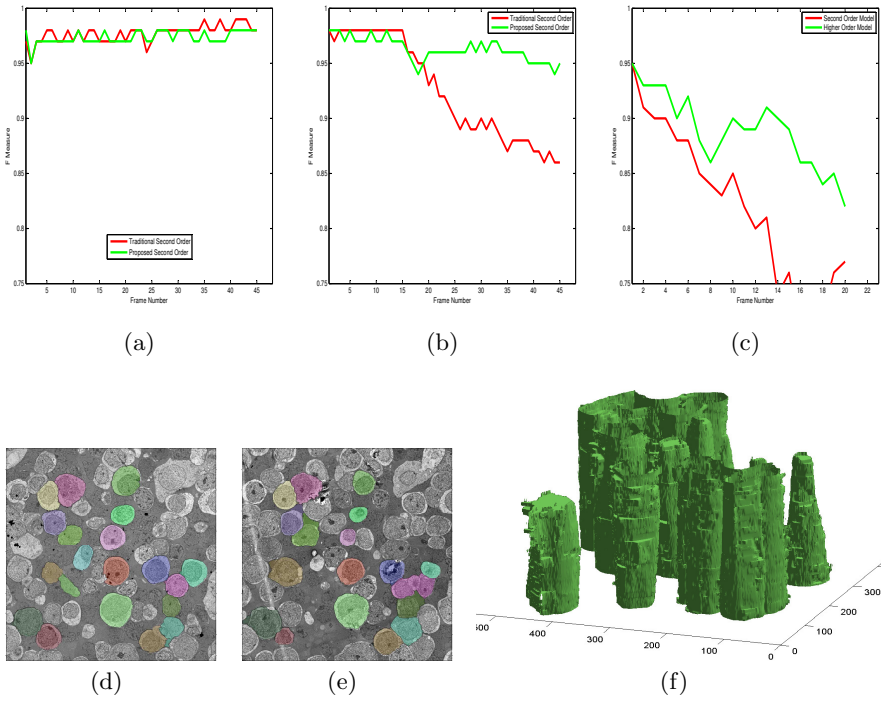


Fig. 3. (Best Viewed in Color) (a) Tracing on simple structure (shown in 1) (b) Performance on medium difficulty structure. (c) Performance on difficult target. (d) and (e) Sample tracing results on different z slices. (f) 3D reconstruction of 20 traced structures.

ple structure tracing in 3D. A total of 30 interacting structures were traced in

parallel. The performance of the algorithm is promising (see Figure 3(f)) with quantitative results in Table 1. Losing the trace of a structure usually happens when there are arbitrary appearance variations (also caused due to illumination artifacts).

Table 1. F measures for multiple structure tracing

#Slices	#Contours	Traditional	Proposed
10	300	0.892	0.906
20	600	0.852	0.872
30	900	0.800	0.823

Conclusions: This paper presented a novel framework for tracing multiple neuronal structures in TEM stacks. Experimental results were presented on data from TEM stacks for single and multiple structure tracing. Future work includes investigation of techniques to scale up tracing to large datasets and testing stability of the algorithm under imperfect initialization.

References

1. Boykov, Y., Kolmogorov, V.: An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. PAMI 26(9), 1124–1137 (2004)
2. Boykov, Y., Veksler, O., Zabih, R.: Fast approximate energy minimization via graph cuts. PAMI 23(11), 1222–1239 (2002)
3. Chklovskii, D., Vitaladevuni, S., Scheffer, L.: Semi-automated reconstruction of neural circuits using electron microscopy. Curr. Op. in Neurobiology (2010)
4. Jain, V., Bollmann, B., Richardson, M., Berger, D., Helmstaedter, M., Briggman, K., Denk, W., Bowden, J.: Boundary learning by optimization with topological constraints. In: CVPR 10. pp. 2488–2495. IEEE (2010)
5. Jurrus, E., Hardy, M., Tasdizen, T., Fletcher, P., Koshevoy, P., Chien, C., Denk, W., Whitaker, R.: Axon tracking in serial block-face scanning electron microscopy. Medical image analysis 13(1), 180–188 (2009)
6. Kaynig, V., Fuchs, T., Buhmann, J.: Neuron geometry extraction by perceptual grouping in sstem images. In: CVPR 10. pp. 2902–2909. IEEE (2010)
7. Kaynig, V., Fuchs, T., Buhmann, J.: Geometrical consistent 3d tracing of neuronal processes in sstem data. In: Jiang, T., Navab, N., Pluim, J., Viergever, M. (eds.) MICCAI 2010, LNCS, vol. 6362, pp. 209–216. Springer Berlin / Heidelberg (2010), http://dx.doi.org/10.1007/978-3-642-15745-5_26, 10.1007/978-3-642-15745-526
8. Kohli, P., Ladickỳ, L., Torr, P.: Robust higher order potentials for enforcing label consistency. IJCV 82(3), 302–324 (2009)
9. Kolmogorov, V., Zabih, R.: What energy functions can be minimizedvia graph cuts? PAMI pp. 147–159 (2004)
10. Leung, T., Malik, J.: Representing and recognizing the visual appearance of materials using three-dimensional textons. IJCV 43(1), 29–44 (2001)