An efficient conditional random field approach for automatic and interactive neuron segmentation

Mustafa Gokhan Uzunbas, Chao Chen*, Dimitris Metaxas

Department of Computer Science, Rutgers University, 110 Frelinghuysen Road, Piscataway, NJ 08854-8019, USA

ABSTRACT

We present a new graphical-model-based method for automatic and interactive segmentation of neuron structures from electron microscopy (EM) images. For automated reconstruction, our learning based model selects a collection of nodes from a hierarchical merging tree as the proposed segmentation. More specifically, this is achieved by training a conditional random field (CRF) whose underlying graph is the watershed merging tree. The maximum a posteriori (MAP) prediction of the CRF is the output segmentation. Our results are comparable to the results of state-of-the-art methods. Furthermore, both the inference and the training are very efficient as the graph is tree-structured.

The problem of neuron segmentation requires extremely high segmentation quality. Therefore, *proofreading*, namely, interactively correcting mistakes of the automatic method, is a necessary module in the pipeline. Based on our efficient tree-structured inference algorithm, we develop an interactive segmentation framework which only selects locations where the model is uncertain for a user to proofread. The uncertainty is measured by the marginals of the graphical model. Only giving a limited number of choices makes the user interaction very efficient. Based on user corrections, our framework modifies the merging tree and thus improves the segmentation globally.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Researchers have been working on mapping of circuit diagrams of nerve systems since 1950s. There is consensus that a detailed map of synaptic connectivity would be very helpful for understanding how the brain causes behavior and how brain malfunctions cause behavioral disorders (Lichtman and Sanes, 2008). With the recent developments in electron microscopy (EM) technology, we have neuronal tissue images of very large scale and high resolution. This empower us to reconstruct neuronal circuit in high quality, to identify different parts such as dendrites, synapses and axons, to better understand *Connectomics* (Sporns et al., 2005), i.e., the structure and functionality of the nervous system (Helmstaedter and Mitra, 2012), and finally, to provide deeper understanding and better diagnostic for neuronal disorders such as Alzheimer's, attention deficit hyperactivity disorder (ADHD), etc.

There are several state-of-the-art EM imaging techniques that can provide nanometer (nm) resolution images. Examples include serial block-face scanning electron microscopy (SBFSEM,

Denk and Horstmann, 2004), serial section transmission electron microscopy (ssTEM, Briggman and Bock, 2012), serial section scanning electron microscopy (ssSEM, Horstmann et al., 2012) and focused-ion-beam based scanning electron microscopy (FIBSEM, Knott et al., 2008). Based on the spatial resolution, we can divide these modalities into *anisotropic* and *isotropic*. The anisotropic ones, ssTEM ($4 \times 4 \times 50$ nm^3, Arganda-Carreras et al., 2012) and ssSEM ($6 \times 6 \times 60$ nm^3, Arganda-Carreras et al., 2013), have a coarse resolution in the third dimension, which should be dealt with separately in the segmentation algorithms. The isotropic or almost isotropic ones, SBF-SEM ($16.5 \times 16.5 \times 23$ nm^3, Kim et al., 2014) and FIBSEM ($10 \times 10 \times 10$ nm^3 and Knott et al., 2008), have (almost) the same resolution in all dimensions. Currently most data are images of the nerve tissue of *Caenorhabditis elegans, Drosophila* or adult mouse cortex.

There has been remarkable progress in developing automatic methods to reconstruct neuronal structures, i.e., decomposing the EM image into separate regions/volumes corresponding to different neuronal cells. The state-of-the-art methods have achieved very high reconstruction quality by effectively using various information such as gradient, texture, intensity and contextual prior knowledge. However, this is not sufficient because of the extremely high requirement from the application. The mislabeling of a single pixel/voxel can cause mistakes such as splitting a single neuron into two or merging two neurons into one. This can lead to catastrophic error in the reconstruction

* Corresponding author. Tel.: +1 848 445 2914.
 E-mail addresses: uzunbas@cs.rutgers.edu (M.G. Uzunbas), chao.chen.cchen@gmail.com (C. Chen), dnm@cs.rutgers.edu (D. Metaxas).

of the topology of the network of neuron structures. Overall, the challenges of automatic segmentation include

- blurred and low contrast membranes,
- complex appearance within cells, and
- high variation in shape of structures (elongated, twisted).

As a necessary complementary step to automatic methods, human experts have to *proofread*, namely, to manually correct the results of automated methods. Thus, interactive segmentation systems which can address large data sets with reduced user interaction have been developed. See Section 2 for a detailed review of automatic and interactive segmentation methods of neuron structures from EM images.

In this paper, we propose a new graphical-model-based framework for the neuron segmentation task. This model leads to not only a principled automatic segmentation algorithm but also an efficient interactive segmentation tool. We propose to construct a conditional random field (CRF) whose underlying graph is the output of the classical *watershed transform* (Beucher and Meyer, 1992). The watershed transform partitions a given image into *superpixels* by simulating a water flooding of the landscape of a given scalar function, e.g. the gradient magnitude or the likelihood of each pixel being the boundary. These over-segmented regions usually form a representation of an image that is much more compact than the original pixel grid. Hence, it is a very common method that is used as preprocessing step of more complex algorithms (Andres et al., 2012a; 2012b; Chklovskii et al., 2010; Jain et al., 2011; Jurrus et al., 2013; Kroeger et al., 2013; Liu et al., 2014; Nunez-Iglesias et al., 2013; Vazquez-Reina et al., 2011). In order to mitigate the over-segmentation effect, one often merges neighboring segments when the minimal function value along the boundary between them (called the *saliency*) is below certain threshold. Considering all saliency thresholds, a hierarchical *merging tree* is constructed (Najman and Schmitt, 1996) in which each leaf node is a segment of the original watershed and each non-leaf node is a merged segment. A *height* function can be assigned to each node according to the minimal saliency threshold at which it disappears (is merged with others). The watershed segmentation at any given threshold can be computed by cutting all tree nodes below

the threshold and taking all leaf nodes of the remaining tree (see Fig. 1(d) and (e)).

We use the watershed merging tree as the underlying graph of our graphical model. The MAP solution of the graphical model assigns labels to the nodes of this hierarchical tree and determines a cut of the tree at different height in different locations. Fig. 1 shows the advantage of our adaptive cut over the classical watershed cut. Three colors correspond to three labels. Black nodes are below the cut. Blue and red nodes are above the cut. Blue nodes are leaf nodes after the cut. Running watershed using a certain threshold usually leads to accurate segments at certain area yet over/under-segmentation at other areas (see Fig. 1(d) and (e)). However, our CRF-based learning algorithm finds a segmentation of higher quality by selecting different saliency thresholds at different locations of the image. Essentially, our algorithm learns from training data how to cut a hierarchical tree adaptively to achieve a better result. See Fig. 1(f). Our automatic segmentation method is not only accurate, but also very efficient. The tree structure of the graphical model allows us to compute exact MAP inference and the marginals very fast. Our method outperforms state-of-the-arts in automatic segmentation of high resolution 2D and 3D isotropic EM images. For 2D experiments, we validate our method on independent slices of ssTEM and ssSEM data. For 3D experiment, we demonstrate on FIBSEM data.

1.1. Interactive segmentation

Proofreading is a necessary yet painful step in neuron segmentation. In a typically framework, a human expert has to check boundaries of neurons one-by-one (see Fig. 6). Based on our CRF model, we develop a novel interactive segmentation framework. Our framework employs the philosophy of active learning and achieves a much higher efficiency in the following two perspectives.

- **User attention:** Instead of the whole volume, we highlight only a few locations for a human expert to proofread. In particular, we highlight locations at which our model has the lowest confidence (cyan and yellow colored pixels in Fig. 5) . This will drastically reduce the time a user needs for each interaction. Furthermore, user
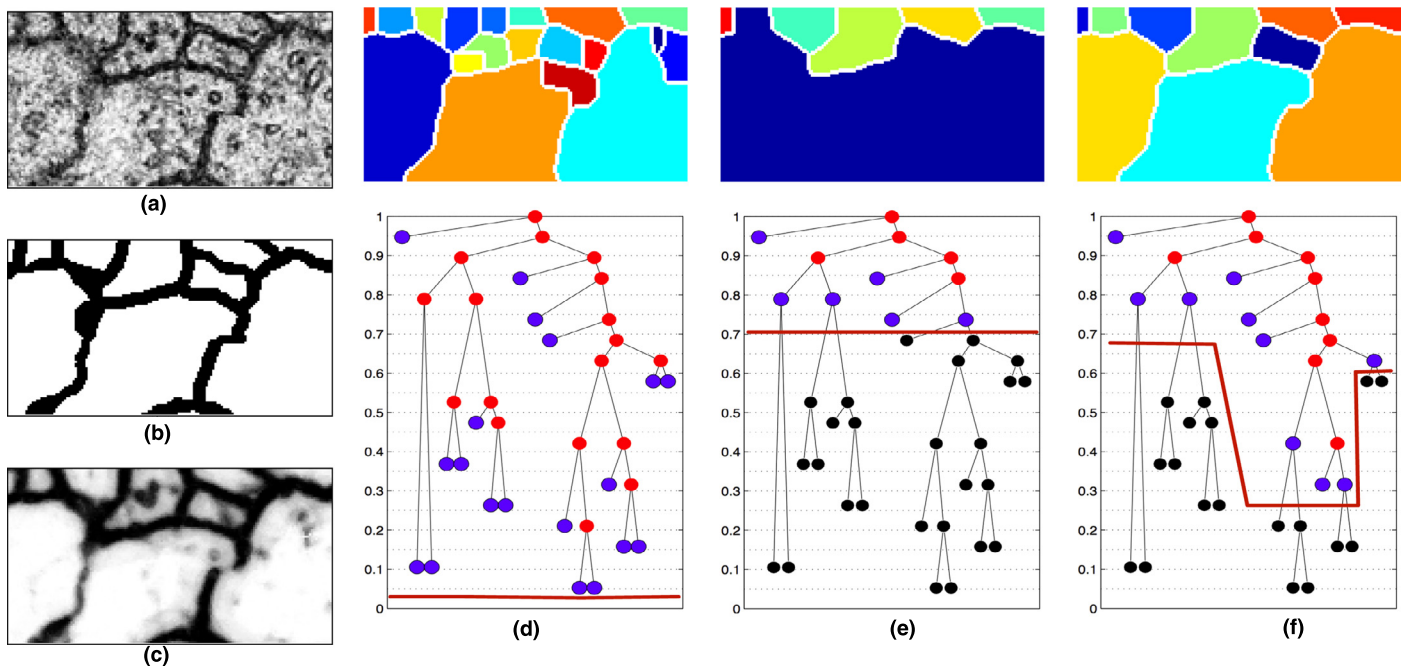

Fig. 1. (a) The EM image patch; (b) the ground truth; (c) the boundary likelihood map (dark pixels have high values); (d) the watershed segmentation and its tree, built using the boundary likelihood map as the landscape function; (e) the watershed segmentation with a higher threshold; (f) the result of our algorithm. (For interpretation of the references to color in the text, the reader is referred to the web version of this article.)