Published in final edited form as:

J Neurosci Methods. 2009 March 30; 178(1): 197–204. doi:10.1016/j.jneumeth.2008.11.008.

Detection of the optimal neuron traces in confocal microscopy images

Zlatko Vasilkoski and Armen Stepanyants

Department of Physics and Center for Interdisciplinary Research on Complex Systems,
Northeastern University, Boston, MA 02115, USA

Abstract

Quantitative methods of analysis of neural circuits rely on large datasets of neurons reconstructed accurately in three dimensions (3D). Due to the complexity of neuronal arbors, large datasets of reconstructed neurons must be generated with automated algorithms. Here, we attempted to automate the process of neuron tracing from sparsely labeled 3D stacks of confocal microscopy images. Our algorithm involves two steps. In the first step, the segmented image of neurites in the stack is voxel-coded. Centers of intensity of consecutively coded wave fronts are connected into a branched structure, which represents a coarse trace of the neurites. In the second step, this trace is optimized with the modified active contour method, which tends to maximize the intensity along the trace while keeping it under tension. To assess the performance of the algorithm we used manual reconstructions of neurons and converted them into artificial stacks of intensity images. These images were traced using the developed algorithm and quantitatively compared to the corresponding manual traces. The optimal traces were on average 6.0% shorter than the manual traces. This reduction in length resulted from the smoothness of the optimal traces, which, in comparison to the manual ones, were built out of shorter segments, and, as a result, were 3.3% less tortuous. The average distance between the optimal and the manual traces was 0.14 μm , and the average distance between their corresponding branch-points was 2.2 μm , illustrating good agreement between the traces.

Introduction

It is virtually impossible to understand how the brain functions without detailed knowledge of neural circuits. Axons and dendrites of neurons can be labeled with a multitude of cell labeling techniques and visualized in three-dimensions (3D) with confocal scanning microscopy (Conchello and Lichtman, 2005). Such images contain valuable information about the connectivity in the brain, despite the facts that only small subsets of neurons are labeled in typical experiments to allow for quantitative analysis, and that synaptic connectivity cannot be resolved reliably on the level of light microscopy. The first restriction can be partially alleviated by new circuit labeling techniques such as multicolor Brainbow (Livet et al., 2007; Lichtman et al., 2008), which allows for labeling of a much larger subsets of neurons through the introduction of different colors. Transsynaptic tracing (Wickersham et al., 2007), which identifies the pre-synaptic partners of a given cell, or array tomography (Micheva and Smith, 2007) can alleviate the second restriction on the visualization of neural circuits.

In spite of the above limitations, quantitative methods of analysis of axonal and dendritic arbors of sparsely labeled neurons have already resulted in numerous findings about synaptic connectivity [see e.g. (Elston and Rosa, 1997; Kozloski et al., 2001; Duan et al., 2002; Stettler et al., 2002; Lubke et al., 2003; Binzegger et al., 2004; Stepanyants et al., 2004; Li et al., 2005; Shepherd et al., 2005; Markram, 2006; Jefferis et al., 2007; Lee and Stevens, 2007; Stepanyants et al., 2008)]. These methods rely on accurate reconstruction of labeled neurons, which currently can only be achieved with manual reconstruction software tools, such as the

ones included in Neurolucida (MicroBrightField Inc.), NeuronJ (Meijering et al., 2004), NeuronStudio (CNIC, Mount Sinai School of Medicine), and FilamentTracer (Bitplane Inc.). However, the manual reconstruction is extremely time consuming and depends on the diligence and the subjective judgment of the user. There is an obvious need to automate the neuron reconstruction process.

Automated reconstruction of neurites from confocal microscopy stacks of images typically contains several of the following steps (Russ, 2007): image restoration (denoising and deconvolution), registration of images within individual image stacks, image segmentation, tracing, feature extraction (cell bodies, spines, boutons, synapses, branch-points), and registration of stacks of images. The purpose of this study is to develop and evaluate the performance of an automated algorithm for neurite tracing. To focus on this aim, we assumed that all the remaining reconstruction steps have been already accomplished or will be performed after the tracing is completed. In particular, we assumed that neurites had been successfully segmented from the background, i.e. they contain no loops or broken segments.

There are generally several classes of algorithms, which can be used to obtain the traces of segmented linear branched structures in 3D. One class of algorithms is based on image thinning methods, which are used for extracting skeletons from binary images [see e.g. (Lee et al., 1994; Palagyi and Kuba, 1998)]. The basic idea behind these methods is an iterative removal of voxels from the surface of the segmented image in a way that preserves the topology of the contained structure. When applied to noisy images of neurites, such algorithms produce voxelated skeletons, which do not follow precisely the intensity in the image and often contain erroneous short branches.

Multi-scale 3D line enhancement filters (Sato et al., 1998; Streekstra and van Pelt, 2002), that are based on the eigenvalues of the Hessian matrix, can facilitate centerline visualization and detection. These filters can extract centerline voxels in non-branched linear structures. Once the centerline is detected, fitting models can be used to determine the most optimal trace of the linear structure. One of the most frequently used models is the active contour model introduced by Kass (Kass et al., 1988; Frangi et al., 1999; Schmitt et al., 2004) that fits the shape of a curve to the skeleton. In most of these models, the user sets the branch-points and they remain fixed during the fitting process.

Another approach is vectorial tracking (Can et al., 1999; Al-Kofahi et al., 2002; Srinivasan et al., 2007; Wang et al., 2007). This class of algorithms exploits local image properties to recursively trace linear structures. The algorithm starts from a seed point, which is typically provided by the user, and tracks the structure of interest by exploring the nearby voxels. The algorithm can be combined with a dynamic real-time updating of the trace direction by the user (Meijering et al., 2004). Vectorial tracking is suitable for semi-automated image analysis of high contrast images. However, this method is not appropriate for precise fully automated tracing of branched structures because the branch- and end-points of the structure have to be specified by the user.

Currently, there are no fully automated tools, which would reliably trace complex neuronal morphology, e.g. axonal arbors of cortical pyramidal neurons. Most methods require from the user to perform pattern recognition tasks, such as identification of branch- and end-points. In this study, we described an automated algorithm for determining the optimal traces of already segmented neurite images. This computerized method of neuron tracing requires no user involvement. It consists of two steps, voxel-coding and optimal tracing, which are described in the first part of this paper. In the second part, we compared the results obtained with the automated tracing procedure to the manual neuron reconstructions.

Materials and Methods

Manual neuron traces

In this study we used dendritic arbors of 20 pyramidal cells reconstructed from layer 3 of monkey superior temporal cortex and prefrontal area 46 (Duan et al., 2002; Duan et al., 2003). These neurons were retrogradely labeled with Lucifer Yellow and reconstructed in 3D with custom-designed morphometry software NeuroZoom (Nimchinsky et al., 1996; Nimchinsky et al., 2000). Neuron reconstructions were obtained from <http://NeuroMorpho.Org> (Ascoli, 2006) in the SWC format (Cannon et al., 1998).

Converting neuron traces into image stacks

For every manually reconstructed cell, we generated a 3D stack of intensity images by using the following procedure. We assumed that the underlying neurite fluorescence, F_N , is distributed uniformly throughout the arbor of the labeled neuron. Since in the present study we are only interested in finding the optimal traces of already segmented images, background level of fluorescence, F_B , was assumed to be zero. To simulate the light scattering in the tissue and in the microscope we linearly blurred the fluorescence of the labeled tissue by convolving it with a point spread function (*PSF*). For convenience, the *PSF* was chosen to be Gaussian with equal standard deviations in all three dimensions, $\sigma_{x,y,z} = 1\mu\text{m}$. As a result of this procedure, we obtained the expected fluxes of photons arriving at the detector from individual locations in the tissue. These fluxes, integrated at the detector over time, result in the expected counts of photons arriving from individual voxels in the image stack, $\mu(x,y,z)$. For simplicity, voxel size was chosen to be $s = 1\mu\text{m}$ in all three dimensions. The actual photon counts, $n(x,y,z)$, were randomly chosen from the Poisson distributions, $P(n/x,y,z)$, with the corresponding expected values, $\mu(x,y,z)$. This procedure can be summarized as:

$$\begin{aligned} PSF(x, y, z) &= \frac{1}{(2\pi)^{3/2} \sigma_x \sigma_y \sigma_z} \exp \left[-\frac{x^2}{2\sigma_x^2} - \frac{y^2}{2\sigma_y^2} - \frac{z^2}{2\sigma_z^2} \right] \\ \mu(x, y, z) &= s^3 \int F(x', y', z') PSF(x - x', y - y', z - z') dx' dy' dz' \\ P(n|x, y, z) &= \frac{\exp[-\mu(x,y,z)] \mu(x,y,z)^n}{n!}. \end{aligned} \quad (1)$$

In this expression, $F(x,y,z)$ is the underlying fluorescence of the tissue. It assumes the values of F_N for points belonging to the interior of the neurites and F_B for the exterior points.

To create an artificial image stack from a manually traced neuron we first voxelated the trace by assigning values of $F_N = 20\mu\text{m}^{-3}$ to all the voxels overlapping with the trace and $F_B = 0\mu\text{m}^{-3}$ to the rest of the voxels. This procedure yielded the function $F(x,y,z)$, which was next convolved (second expression in Eq. 1) with the *PSF* to produce the expected photon counts in the image stack, $\mu(x,y,z)$. Next, the actual photon counts, $n(x,y,z)$, were randomly generated for each voxel from the corresponding Poisson distributions (last expression in Eq. 1). In order to avoid loops, voxels with zero actual photon counts in the interior of neurites were assigned the photon counts of 1. Finally, the intensity $I(x,y,z)$ in the image stack was obtained by scaling down the actual photon counts to the [0 1] interval. The value of $F_N = 20\mu\text{m}^{-3}$ was chosen because it produced image stacks, where the intensity along the neurites was similar to or more variable than that in typical confocal microscopy images of neurites (Figure 1). Our method is not sensitive to particular values of this parameter.

The optimal trace

We begin the description of the method for finding the optimal traces of linear branched structures in confocal microscopy images by first looking at a simplified example. Consider a 3D continuous intensity image, $I(\vec{r})$, of a non-branched line with known start- and end-points,

$\vec{r}_i$ and $\vec{r}_f$. One can employ active contour models (Kass et al., 1988) to obtain the optimal trace of this image. To this end, a fitness function has to be constructed which favors smooth traces connecting points $\vec{r}_i$ and $\vec{r}_f$ in a way that maximizes the intensity along the trace.

We used a parametric representation of the trace line, $\vec{r}(t)$, $t \in [0,1]$, which originates at the point $\vec{r}(0) = \vec{r}_i$ and terminates at the point $\vec{r}(1) = \vec{r}_f$. On the one hand, the optimal trace has to pass through the high intensity places in the image and on the other hand, the trace has to be kept under tension to prevent it from zigzagging from one high intensity voxel to another. The trade-off between these two requirements can be captured by the maximization of the fitness functions containing some of the following terms (Kass et al., 1988; Frangi et al., 1999; Schmitt et al., 2004),

$$F = a \int_0^1 I(\vec{r}(t)) dt + b \int_0^1 I(\vec{r}(t)) \left| \frac{d\vec{r}(t)}{dt} \right| dt - c \int_0^1 \left| \frac{d\vec{r}(t)}{dt} \right| dt - d \int_0^1 \left| \frac{d\vec{r}(t)}{dt} \right|^2 dt. \quad (2)$$

Other terms in the fitness functions had also been considered in the past (Kass et al., 1988; Frangi et al., 1999).

The first two terms in this fitness function benefit from routing the trace through the high intensity regions. The first term gives the average intensity along the trace, while the second corresponds to the integrated intensity along the trace. The third and fourth terms were included to keep the trace under tension. The third term corresponds to the length of the trace, and the fourth term can be interpreted as the elastic energy due to stretching of the trace. Adjustable coefficients $a, b, c, d > 0$ capture the relative contributions of the four terms to the overall fitness, F . It was empirically determined that it is sufficient to include any two terms, where one is related to intensity and the other is related to tension. Here, in order to simplify the description of the methods, we considered the fitness function, which consists of the first and the last terms of Eq. 2. All the results presented in this study were obtained by maximizing this fitness function.

Since confocal microscopy stacks of images are voxellated and not continuous, it is more convenient to rewrite the fitness function in a discrete notation. To do this, we approximated the trace with a broken line consisting of $N-1$ straight segments connecting N vertices, and replaced the integration with summation,

$$F(\{\vec{r}_k\}) = \frac{1}{\lambda} \sum_{k=1}^N I(\vec{r}_k) - \alpha \lambda \sum_{k=1}^{N-1} |\vec{r}_{k+1} - \vec{r}_k|^2. \quad (3)$$

Coefficients a and d of Eq. 2 were replaced with $1/\lambda$ and $\alpha\lambda$, where λ denotes the average density of segments along the trace (number of segments per unit length of the trace), and the dimensionless parameter $\alpha > 0$ controls the tension in the trace. The fitness function in Eq. 3 scales linearly with the overall length of the trace. The parameter λ in this equation was introduced to make the two terms of the fitness function independent on the segment density, which is an adjustable parameter.

To make use of the above expression, we need to be able to estimate the intensity in the image stack at the positions specified by vertices, $\vec{r}_k$ (the first term in the Eq. 3). To this end, intensity in the image stack is interpolated onto every position $\vec{r}_k$ by using a Gaussian interpolation procedure (Stepanyants et al., 2008). This results in,

$$F(\vec{r}_k) = \frac{1}{\lambda} \sum_{k=1}^N \sum_i I(\vec{R}_i) s^3 \frac{e^{-\left(\vec{r}_k - \vec{R}_i\right)^2 / 2\sigma^2}}{(2\pi)^{3/2} \sigma^3} - \alpha \lambda \sum_{k=1}^{N-1} \left| \vec{r}_{k+1} - \vec{r}_k \right|^2. \quad (4)$$

Parameter s in this equation denotes the voxel size and vector $\vec{R}_i$ specifies the position of the center of the voxel i . Summation over index i in the first term of Eq. 4 runs over all the voxels in the image stack. In practice, due to a fast decay of the Gaussian function ($\sigma \approx 1 \mu\text{m}$), this sum can be restricted to a small number of voxels in the vicinity of the point $\vec{r}_k$.

If the layout of the initial trace is sufficiently close to that of the optimal solution, the optimization problem of Eq. 4 can be solved with an iterative procedure referred to as the gradient ascent algorithm (Boyd and Vandenberghe, 2004):

$$\begin{aligned} \vec{r}_k^{n+1} &= \vec{r}_k^n + \beta \left. \frac{\partial F(\vec{r}_k)}{\partial \vec{r}_k} \right|_{\vec{r}_k = \vec{r}_k^n} = \\ &= \vec{r}_k^n + \beta \left[-\frac{1}{\lambda} \sum_i \left(\vec{r}_k^n - \vec{R}_i \right) I(\vec{R}_i) s^3 \frac{e^{-\left(\vec{r}_k^n - \vec{R}_i\right)^2 / 2\sigma^2}}{(2\pi)^{3/2} \sigma^5} - 2\alpha \lambda \left(2\vec{r}_k^n - \vec{r}_{k+1}^n - \vec{r}_{k-1}^n \right) \right] \end{aligned} \quad (5)$$

In this expression, $\vec{r}_k^n$ denotes the position of the vertex k of the trace at an iteration step n of the gradient ascent algorithm. As the start- and end-points of the trace are fixed, index k in Eq. 5 runs from 2 to $N-1$. Parameter $\beta > 0$ controls the size of the iteration step.

The gradient ascent procedure in Eq. 5 is easily generalized on the case of a branched structure,

$$\vec{r}_k^{n+1} = \vec{r}_k^n + \beta \left[-\frac{1}{\lambda} \sum_i \left(\vec{r}_k^n - \vec{R}_i \right) I(\vec{R}_i) s^3 \frac{e^{-\left(\vec{r}_k^n - \vec{R}_i\right)^2 / 2\sigma^2}}{(2\pi)^{3/2} \sigma^5} - 2\alpha \lambda \sum_l \left(\vec{r}_k^n - \vec{r}_{k_l}^n \right) \right]. \quad (6)$$

Here, index k_l enumerates all the segments which are connected to the vertex k . If vertex k is an intermediate-point, the second sum in Eq. 6 contains only two terms, and Eq. 6 reduces to Eq. 5. For a bifurcation point, the second sum will contain three terms, for a trifurcation – four, etc. No calculations are needed for the start- and end-points of the branched structure, as these points must remain stationary.

For small values of parameter α in Eq. 6, the trace becomes very stretchy and the optimal solution can double on itself at high intensity regions in the image stack. This can happen if the fitness function, Eq. 4, benefits more from threading the trace through the same high intensity region several times than it is penalized for the increase in the length of the trace. To prevent such solutions of the optimization problem, the trace must be sufficiently stiff. Analysis of Eq. 4 resulted in the following lower bound for the parameter α ,

$$\alpha_{\min} = \max_{\vec{R}} \left[\sum_i I(\vec{R}_i) s^3 \frac{e^{-\left(\vec{R} - \vec{R}_i\right)^2 / 2\sigma^2}}{(2\pi)^{3/2} \sigma^3} \right]. \quad (7)$$

The maximum in this expression is taken over all the positions $\vec{R}$ on the segmented image. The value of $\alpha_{\min}$ depends on the distribution of intensity in the image stack, the voxel size s , and

the parameter σ . In practice, for images of neurites with intensity in the $[0\ 1]$ range, the value of $\alpha_{\min}$ is roughly 0.1–0.3.

Very large values of the parameter α may result in traces which are excessively straight and do not faithfully follow the intensity in the image in places where neurites make sharp turns. For some value of this parameter, $\alpha_{\max}$, the algorithm will trace the most tortuous parts of the arbor along a straight line, rather than following the image intensity. Hence, the numerical value of $\alpha_{\max}$ is dependent on the type of the imaged neurites. For the dataset of images used in this study, it was empirically established that $\alpha_{\max} = 1.5$.

Since parameter β controls the step of the gradient ascent algorithm, increasing the value of this parameter would lead to a faster convergence of the trace to the optimal solution. However, large values of this parameter may result in unstable or diverging traces. Stability analysis of the above iterative procedure places an upper bound on the possible values of the step size at $\beta_{\max} = 1/(4\alpha\lambda)$.

To further improve the stability of the trace, long segments were subdivided and short segments were combined at every iteration step of the gradient ascent algorithm. The following values of the parameters were used in this study: $s = 1\ \mu\text{m}$, $\sigma = 1\ \mu\text{m}$, $\lambda = 1.0\ \mu\text{m}^{-1}$, $\alpha = 0.3$, $\beta = \beta_{\max}/2 = 0.42\ \mu\text{m}$. These numerical parameters led to stable convergence of the trace to the optimal solution. The fitness function, Eq. 4, increased monotonically throughout the iterations and the procedure was deemed convergent to the optimal solution after 40 iteration steps.

Results

To evaluate the performance of the algorithm, we converted the dendritic arbors of 20 pyramidal cells, which were manually reconstructed from layer 3 of monkey superior temporal cortex and prefrontal area 46 (Duan et al., 2002; Duan et al., 2003), into artificial stacks of images (Figure 1). This was done by blurring the manual traces and introducing Poisson noise to the images in order to imitate low intensity imaging (see Materials and Methods section for the details). Voxel size in the artificial image stacks was chosen to be $1\ \mu\text{m}^3$. Creating the artificial stacks of images from manually traced neurons, rather than using real confocal image stacks, was done for two reasons. First, we wanted to be able to compare the results of our algorithm to both the intensity images and the manual traces. Second, we wanted to avoid the segmentation problem, which is beyond the scope of this study. For the same reason, the background intensity in the artificial image stacks was set to zero, and we only considered cells which did not lead to the formation of loops in the process of conversion.

Below, we describe an automated tracing procedure, which should be performed following the segmentation of neurites. Namely, we assumed that voxels of interest had been successfully separated from their background, as shown in Figure 1A. The segmented voxels need to be further analyzed to extract the optimal 3D trace of the neurites. This was done by using the following two steps: voxel-coding and optimal tracing.

The voxel-coding and optimal tracing

A simple way to obtain the topology of a 3D segmented structure and to examine its connectedness is the voxel-coding algorithm (Zhou et al., 1998; Zhou and Toga, 1999; Shikata et al., 2004; Yamasaki et al., 2006). In this recursive algorithm, a label (a natural number) is associated with each voxel of the segmented image (Figure 2A). For each connected region in the image stack, the voxel-coding algorithm was initiated by selecting a seed voxel, which in principle can be randomly chosen. The seed voxel was assigned the label of 1. At the next step, the voxel-coding algorithm finds all the voxels that are neighboring to the seed voxel. The neighborhood of a voxel is specified by a 3D structural element, which in this study was the

26 nearest neighbors. This neighborhood was referred to as the wave front of the seed voxel, and label 2 was assigned to all of its voxels. At each step, a new wave front was created and consecutively labeled. This encoding procedure was repeated until all of the segmented voxels in the image were labeled.

As the wave propagated through the branch-points, voxel-coding produced wave fronts consisting of two or more disconnected regions. For every disconnected wave front region the center of intensity (center of mass) was obtained. These points formed the wave trace (Figure 2B). The end-points in the wave trace corresponded to the centers of intensity of the terminal wave front regions. The branch-points were defined as the centers of intensity of the front regions, which divided at the subsequent steps. The remaining wave trace points were termed intermediate. The coordinates of the end-, branch-, and intermediate-points, and their topological structure was conveniently recorded in the SWC file format (Cannon et al., 1998). In this study, there was only a single segmented region in each image stack (the entire dendritic arbor), and the voxel-coding algorithm was initialized at the soma of each cell. In general, the described procedure must be applied to all the disconnected neurites present in the segmented image stack.

Because the voxel-coding procedure was performed on a binary segmented image, the resulting wave trace did not follow precisely the intensity in the image stack. The wave trace was not smooth and the branch-points were placed imprecisely (Figure 2B). Next, the optimal tracing procedure was applied to correct for these artifacts (see Materials and Methods). This procedure was based on the modified active contour method (Kass et al., 1988; Frangi et al., 1999; Schmitt et al., 2004), and it used the wave trace as a starting layout. The optimal tracing procedure iteratively moved the intermediate- and branch-points of the trace and converged towards the optimal structure within few tens of iterative steps, while keeping the end-points stationary.

Comparison of the optimal and the manual traces

First, we obtained the optimal traces for the 20 artificially created image stacks and visually examined how well these traces fit the intensity images (Figure 3). Figures 3B and 3C illustrate that the optimal trace provides a reasonably good 3D fit to the intensity image. Second, we quantitatively compared the optimal traces to the corresponding manual reconstructions. For this comparison, we selected five features of neuron morphology: the total arbor length, the tortuosity of neuronal branches, the number of end-points, the proximity of the arbors, and the placement of branch-points.

Figure 4A shows the comparison of the total arbor lengths between the manual and the optimal traces. The 20 data points represent individual analyzed cells. The length of the optimal trace was highly correlated (adjusted $R^2 = 0.9995$) with the length of the manual trace, and was on average $6.0 \pm 0.9\%$ (mean $\pm$ SD) shorter. This difference resulted from the fact that manual neuron traces are composed of straight segments of variable lengths, which are typically several micrometers long. The optimal trace was formed by shorter segments (1 μm on average), and it provided a smoother trace of the intensity image, resulting in a shorter length.

This point is reflected in the comparison of the optimal and the manual trace tortuosities (Stepanyants et al., 2004). Tortuosity index, T , of a dendritic branch segment is defined as the ratio of the segment length, L , to the geometric distance between its end-points, R (see the inset of Figure 4B). We calculated the tortuosities for non-branching segments of different lengths, for all the 20 traced neurons. Segments containing branch-points were excluded from this analysis, as their tortuosity indices are influenced by branching. The results are shown in Figure 4B where the tortuosity indices of the manual and the optimal traces are plotted for segments of different lengths. As was expected, the tortuosity index is an increasing function of the

segment length (Stepanyants et al., 2004). The optimal trace tortuosity index was roughly 3.3% smaller than that for the manual trace, indicating a smoother fit to the intensity image.

The voxel-coding procedure is sensitive to the smoothness of the segmented image. Because this procedure is applied to binary images of neurites, relatively large (few voxels in size) irregularities on the surface of the neurites may result in creation of short erroneous terminal branches. Similarly, voxel-coding can omit some of the existing terminal branches that are short. In this study, these branches were only few voxels (or micrometers) in length. In Figure 4C we compared the numbers of end-points in the manual and the optimal traces. It can be seen that the numbers of end-points in the manual and the optimal traces are highly correlated (adjusted $R^2 = 0.99$), and that for many of the analyzed cells these numbers are the same. However, for some cells the difference in the numbers of terminal branches was as high as 2. The average absolute difference, 0.85 terminal branches per cell, was small in comparison with the total number of terminal branches per cell, 42 ± 10 . As the erroneous terminal branches and the branches that were missed by the algorithm were only few micrometers in length, they have little impact on the overall cell morphology.

To further quantify the accuracy of the automated tracing procedure, we defined the measure of proximity between two traces. To this end, we first re-segmented both traces into short segments (shorter than $0.25 \mu\text{m}$) and determined the centers of these segments. Second, for every segment center on the first trace we found the corresponding closest segment center on the second trace and determined the Euclidian distance between the two. Third, a weighted average of these distances was calculated by using the products of the corresponding segment lengths as weights. Fourth, the traces were switched around and the last two steps were repeated producing an alternative weighted average. Finally, the two weighted distances were averaged resulting in the distance measure that is symmetric with respect to both traces. According to this measure, the average distance between the segments in the manual and the wave traces was $0.31 \mu\text{m}$, and 95% of the corresponding segment centers were separated with distances in the $0.005 \mu\text{m} - 2.03 \mu\text{m}$ range (95% confidence interval). In comparison, the average distance between the segments in the manual and the optimal traces was much lower, $0.14 \mu\text{m}$ ($0.004 \mu\text{m} - 0.63 \mu\text{m}$ 95% confidence interval), emphasizing the importance of the optimal tracing procedure.

As the final comparison between the manual and the optimal traces, we examined the accuracy of branch-point placement. The average distance between the corresponding branch-points in the manual and the optimal traces was $2.2 \pm 1.3 \mu\text{m}$. Since the size of a typical bifurcation region in our images was more than $10 \mu\text{m}$ in diameter, the branch point placement was deemed sufficiently accurate. As was expected (Figure 2B), the average distance between the corresponding branch-points in the manual and the wave traces was much larger, $4.0 \pm 2.0 \mu\text{m}$, again illustrating the benefits of the optimal tracing procedure.

Discussion

Algorithms for automated neuron reconstruction are essential for realistic modeling of neural circuits. In recent years, there has been a dramatic increase in the number of studies, which rely on large sets of high quality neuron reconstructions [see (Stepanyants and Chklovskii, 2005; Ascoli, 2007) for review], and this trend will continue to grow with the advances in experimental techniques and computational capabilities.

In this paper, we described an algorithm that could be used as part of an automated procedure for reconstruction of neural circuits imaged with confocal microscopy. The algorithm was aimed at automating the process of detection of the optimal traces from already segmented neurite images. Segmenting neurites from their background can be a very difficult problem

when the image quality is low or the density of labeled neurites is high. Poor labeling or image quality can lead to low contrast and beaded appearance of neurites. Segmentation of such images often results in broken structures that need to be reconnected afterward. Densely labeled neurites may result in the appearance of loops in the segmented image. Such loops will have to be subsequently resolved, which presents a substantial challenge.

As a starting point of our method, we created segmented stacks of images and focused on the automated and accurate neuron tracing. To this end, we blurred 20 manually traced dendritic arbors of pyramidal cells and transformed them into segmented 3D stacks of intensity images. The automated tracing algorithm presented in this study consists of two parts. First, we voxel-coded the segmented intensity image (Zhou et al., 1998; Zhou and Toga, 1999; Shikata et al., 2004; Yamasaki et al., 2006) and extracted the topology of the 3D neuronal structure. Second, to find the optimal trace of the segmented neurites, we smoothed and optimized this structure by using a modified active contour method (Kass et al., 1988; Frangi et al., 1999; Schmitt et al., 2004).

To test this algorithm, we obtained the optimal traces of the 20 artificially created image stacks and compared them to the initial, manual traces. The comparison showed no major differences between the traces, with the average distance between the manual and the optimal traces being only 0.14 μm . The optimal traces were on average 6.0% shorter than the manual ones. This was a consequence of the fact that in comparison to typical manual neuron traces, where the segment lengths often exceed 5–10 μm , the optimal traces in this study were made of shorter segments. As a result, they provided a smoother representation of the intensity image. This point is reflected in the fact that optimal traces were 3.3% less tortuous than the manual ones.

As mentioned in the results section, the voxel-coding algorithm, due to its binary nature, did not faithfully reproduce the number of terminal branches. We quantified the extent of this artifact by comparing the numbers of end-points in the tree structures of the manual and the optimal traces. The discrepancy was on average 0.85 terminal branches per cell. New erroneous branches or branches that were missed by the algorithm were only few micrometers in length and have very little impact on the morphological characteristics of the cells and on synaptic connectivity. Once the optimal trace of a neuron is detected, it becomes feasible to eliminate the erroneous branches and to find the missed branches by searching in the volume surrounding the optimal trace.

While detecting the optimal neuronal traces from the stacks of intensity images we made no attempt to render the shapes of individual neuronal branches in 3D (Schmitt et al., 2004). The reason for this is the fact that the thickness of axons and dendrites of most neural classes in the cerebral cortex is below the resolution of confocal microscopy. For example, the average (by length) radius of the dendritic branches of cortical pyramidal cells in the mouse neocortex is 0.45 μm , while the average radius of axonal branches is roughly 3 times smaller (Braitenberg and Schüz, 1998). Hence, the details of neuronal morphology related to the thickness of branches cannot be resolved reliably with confocal microscopy. These types of analyses are better addressed with electron (Fiala, 2005; Briggman and Denk, 2006) or high-resolution light microscopy (Micheva and Smith, 2007; Cai et al., 2008).

The ultimate goal of this study is to facilitate a complete reconstruction of sparsely labeled subsets of neural circuits on a large scale, e.g. on the scale of the entire mouse brain. Such reconstruction will require fast and accurate processing of large numbers of confocal image stacks. All calculations in this study were performed on a 3.6 GHz computer and required approximately 2 minutes of processing time per cell, which was divided evenly between the voxel-coding and the optimal tracing parts of the algorithm. In addition, comparison of the

manual and the optimal traces showed that the algorithm is sufficiently accurate, which makes it potentially useful for the large-scale neural circuit reconstruction projects.

Acknowledgments

Authors thank Tarec Fares from Northeastern University for discussions related to the subject of this study and for insightful comments on the manuscript. This work was supported by the NIH grant NS063494.

References

- Al-Kofahi KA, Lasek S, Szarowski DH, Pace CJ, Nagy G, Turner JN, Roysam B. Rapid automated three-dimensional tracing of neurons from confocal image stacks. *IEEE Trans Inf Technol Biomed* 2002;6:171–187. [PubMed: 12075671]
- Ascoli GA. Mobilizing the base of neuroscience data: the case of neuronal morphologies. *Nat Rev Neurosci* 2006;7:318–324. [PubMed: 16552417]
- Ascoli GA. Successes and rewards in sharing digital reconstructions of neuronal morphology. *Neuroinformatics* 2007;5:154–160. [PubMed: 17917126]
- Binzegger T, Douglas RJ, Martin KA. A quantitative map of the circuit of cat primary visual cortex. *J Neurosci* 2004;24:8441–8453. [PubMed: 15456817]
- Boyd, SP.; Vandenberghe, L. *Convex optimization*. Cambridge University Press; Cambridge, UK; New York: 2004.
- Braitenberg, V.; Schüz, A. *Cortex: statistics and geometry of neuronal connectivity*. Vol. 2nd thoroughly rev. Edition. Springer; Berlin; New York: 1998.
- Briggman KL, Denk W. Towards neural circuit reconstruction with volume electron microscopy techniques. *Curr Opin Neurobiol* 2006;16:562–570. [PubMed: 16962767]
- Cai H, Xu X, Lu J, Lichtman J, Yung SP, Wong ST. Using nonlinear diffusion and mean shift to detect and connect cross-sections of axons in 3D optical microscopy images. *Med Image Anal*. 2008
- Can A, Shen H, Turner JN, Tanenbaum HL, Roysam B. Rapid automated tracing and feature extraction from retinal fundus images using direct exploratory algorithms. *IEEE Trans Inf Technol Biomed* 1999;3:125–138. [PubMed: 10719494]
- Cannon RC, Turner DA, Pyapali GK, Wheal HV. An on-line archive of reconstructed hippocampal neurons. *J Neurosci Methods* 1998;84:49–54. [PubMed: 9821633]
- Conchello JA, Lichtman JW. Optical sectioning microscopy. *Nat Methods* 2005;2:920–931. [PubMed: 16299477]
- Duan H, Wearne SL, Morrison JH, Hof PR. Quantitative analysis of the dendritic morphology of corticocortical projection neurons in the macaque monkey association cortex. *Neuroscience* 2002;114:349–359. [PubMed: 12204204]
- Duan H, Wearne SL, Rocher AB, Macedo A, Morrison JH, Hof PR. Age-related dendritic and spine changes in corticocortically projecting neurons in macaque monkeys. *Cereb Cortex* 2003;13:950–961. [PubMed: 12902394]
- Elston GN, Rosa MG. The occipitoparietal pathway of the macaque monkey: comparison of pyramidal cell morphology in layer III of functionally related cortical visual areas. *Cereb Cortex* 1997;7:432–452. [PubMed: 9261573]
- Fiala JC. Reconstruct: a free editor for serial section microscopy. *J Microsc* 2005;218:52–61. [PubMed: 15817063]
- Frangi AF, Niessen WJ, Hoogeveen RM, van Walsum T, Viergever MA. Model-based quantitation of 3-D magnetic resonance angiographic images. *Ieee Transactions on Medical Imaging* 1999;18:946–956. [PubMed: 10628954]
- Jefferis GS, Potter CJ, Chan AM, Marin EC, Rohlfsing T, Maurer CR Jr, Luo L. Comprehensive maps of *Drosophila* higher olfactory centers: spatially segregated fruit and pheromone representation. *Cell* 2007;128:1187–1203. [PubMed: 17382886]
- Kass M, Witkin A, Terzopoulos D. Snakes: Active contour models. *International Journal of Computer Vision* 1988;1:321–331.

- Kozloski J, Hamzei-Sichani F, Yuste R. Stereotyped position of local synaptic targets in neocortex. *Science* 2001;293:868–872. [PubMed: 11486089]
- Lee S, Stevens CF. General design principle for scalable neural circuits in a vertebrate retina. *Proc Natl Acad Sci U S A* 2007;104:12931–12935. [PubMed: 17646664]
- Lee TC, Kashyap RL, Chu CN. Building skeleton models via (3-D) medial surface/axis thinning algorithms. *CVGIP: Graph Models Image Process* 1994;56:462–478.
- Li Y, Brewer D, Burke RE, Ascoli GA. Developmental changes in spinal motoneuron dendrites in neonatal mice. *J Comp Neurol* 2005;483:304–317. [PubMed: 15682391]
- Lichtman JW, Livet J, Sanes JR. A technicolour approach to the connectome. *Nat Rev Neurosci* 2008;9:417–422. [PubMed: 18446160]
- Livet J, Weissman TA, Kang H, Draft RW, Lu J, Bennis RA, Sanes JR, Lichtman JW. Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. *Nature* 2007;450:56–62. [PubMed: 17972876]
- Lubke J, Roth A, Feldmeyer D, Sakmann B. Morphometric analysis of the columnar innervation domain of neurons connecting layer 4 and layer 2/3 of juvenile rat barrel cortex. *Cereb Cortex* 2003;13:1051–1063. [PubMed: 12967922]
- Markram H. The blue brain project. *Nat Rev Neurosci* 2006;7:153–160. [PubMed: 16429124]
- Meijering E, Jacob M, Sarria JC, Steiner P, Hirling H, Unser M. Design and validation of a tool for neurite tracing and analysis in fluorescence microscopy images. *Cytometry A* 2004;58:167–176. [PubMed: 15057970]
- Micheva KD, Smith SJ. Array tomography: a new tool for imaging the molecular architecture and ultrastructure of neural circuits. *Neuron* 2007;55:25–36. [PubMed: 17610815]
- Nimchinsky EA, Hof PR, Young WG, Morrison JH. Neurochemical, morphologic, and laminar characterization of cortical projection neurons in the cingulate motor areas of the macaque monkey. *J Comp Neurol* 1996;374:136–160. [PubMed: 8891952]
- Nimchinsky EA, Young WG, Yeung G, Shah RA, Gordon JW, Bloom FE, Morrison JH, Hof PR. Differential vulnerability of oculomotor, facial, and hypoglossal nuclei in G86R superoxide dismutase transgenic mice. *J Comp Neurol* 2000;416:112–125. [PubMed: 10578106]
- Palagyi K, Kuba A. A 3D 6-subiteration thinning algorithm for extracting medial lines. *Pattern Recognition Letters* 1998;19:613–627.
- Russ, JC. Vol. 5th Edition. CRC/Taylor and Francis; Boca Raton: 2007. The image processing handbook.
- Sato Y, Nakajima S, Shiraga N, Atsumi H, Yoshida S, Koller T, Gerig G, Kikinis R. Three-dimensional multi-scale line filter for segmentation and visualization of curvilinear structures in medical images. *Med Image Anal* 1998;2:143–168. [PubMed: 10646760]
- Schmitt S, Evers JF, Duch C, Scholz M, Obermayer K. New methods for the computer-assisted 3-D reconstruction of neurons from confocal image stacks. *Neuroimage* 2004;23:1283–1298. [PubMed: 15589093]
- Shepherd GM, Stepanyants A, Bureau I, Chklovskii D, Svoboda K. Geometric and functional organization of cortical circuits. *Nat Neurosci* 2005;8:782–790. [PubMed: 15880111]
- Shikata H, Yin Z, Park Y, Kitaoka H, Sato Y, Johkoh T, Nakamura H, Tamura S. An algorithm for localizing branch points of pulmonary vessels for nonrigid registration of the lungs. *Systems and Computers in Japan* 2004;35:1613–1623.
- Srinivasan R, Zhou X, Miller E, Lu J, Litchman J, Wong ST. Automated axon tracking of 3D confocal laser scanning microscopy images using guided probabilistic region merging. *Neuroinformatics* 2007;5:189–203. [PubMed: 17917130]
- Stepanyants A, Chklovskii DB. Neurogeometry and potential synaptic connectivity. *Trends Neurosci* 2005;28:387–394. [PubMed: 15935485]
- Stepanyants A, Tamás G, Chklovskii DB. Class-specific features of neuronal wiring. *Neuron* 2004;43:251–259. [PubMed: 15260960]
- Stepanyants A, Hirsch JA, Martinez LM, Kisvárdy ZF, Ferecskó AS, Chklovskii DB. Local potential connectivity in cat primary visual cortex. *Cereb Cortex* 2008;18:13–28. [PubMed: 17420172]
- Stettler DD, Das A, Bennett J, Gilbert CD. Lateral connectivity and contextual interactions in macaque primary visual cortex. *Neuron* 2002;36:739–750. [PubMed: 12441061]

- Streekstra GJ, van Pelt J. Analysis of tubular structures in three-dimensional confocal images. *Network* 2002;13:381–395. [PubMed: 12222820]
- Wang, J.; Zhou, X.; Lu, J.; Lichtman, J.; Chang, S-F.; Wong, STC. IEEE, International Symposium of Biomedical Imaging (ISBI). Washington D. C.: 2007. Dynamic local tracing for 3D axon curvilinear structure detection from microscopic image stack..
- Wickersham IR, Lyon DC, Barnard RJ, Mori T, Finke S, Conzelmann KK, Young JA, Callaway EM. Monosynaptic restriction of transsynaptic tracing from single, genetically targeted neurons. *Neuron* 2007;53:639–647. [PubMed: 17329205]
- Yamasaki T, Isokawa T, Matsui N, Ikeno H, Kanzaki R. Reconstruction and simulation for three-dimensional morphological structure of insect neurons. *Neurocomputing* 2006;69:1043–1047.
- Zhou Y, Toga AW. Efficient Skeletonization of Volumetric Objects. *IEEE Transactions on Visualization and Computer Graphics* 1999;05:196–209.
- Zhou Y, Kaufman A, Toga AW. Three-dimensional skeleton and centerline generation based on an approximate minimum distance field. *Visual Computer* 1998;14:303–314.

Figure 1. Artificially created image stacks. **A.** Maximum intensity projection of an image stack, which was artificially created from a manually traced pyramidal cell dendrite. **B.** yz and **C.** xy maximum intensity projections of the selected region from **A**. The manual trace is shown in green. Scale bars are $50\ \mu\text{m}$ in **A** and $10\ \mu\text{m}$ in **B** and **C**.

Figure 2.

Illustration of the voxel-coding method. **A.** Maximum intensity projection of a small part of a dendritic arbor. A 3D wave is initiated at an arbitrary voxel in the image (yellow circle). At subsequent steps, the new wave fronts are determined based on the choice of the structural element ($3 \times 3 \times 3$ cube in the figure). The branch-point (green asterisk) is defined as the center of intensity of the wave front that splits in two at the subsequent step. The end-points (blue asterisks) correspond to the centers of intensity of the terminal wave fronts. Intermediate-points are the centers of intensity of the remaining wave fronts. **B.** Subsequent end-, branch-, and intermediate-points are connected with straight segments to form the wave trace. This trace (yellow line) is shown superimposed on the maximum intensity projected image from Figure 1B. Scale bars is $10 \mu\text{m}$.

Figure 3.

The optimal trace. **A.** The optimal trace (red line) is shown superimposed onto the maximum intensity projected image from Figure 1. **B.** yz and **C.** xy maximum intensity projections of the selected region from A. Scale bars are 50 μm in A and 10 μm in B and C.

Figure 4.

Comparison of the optimal and the manual neuron traces. **A.** Comparison of lengths of the optimal and the manual traces. Individual points represent different cells. **B.** Comparison of tortuosities of the optimal and the manual traces. Tortuosity index, T , is calculated for non-branching dendritic segments of different lengths as the ratio of the geometric distance between the segment end-points, R , and the segment length, L (inset). **C.** Comparison of the numbers of end-points in the optimal and the manual traces.