Neuron Tracing in Perspective

Erik Meijering


ABSTRACT

The study of the structure and function of neuronal cells and networks is of crucial importance in the endeavor to understand how the brain works. A key component in this process is the extraction of neuronal morphology from microscopic imaging data. In the past four decades, many computational methods and tools have been developed for digital reconstruction of neurons from images, with limited success. As witnessed by the growing body of literature on the subject, as well as the organization of challenging competitions in the field, the quest for a robust and fully automated system of more general applicability still continues. The aim of this article is to contribute to this by surveying recent developments in the field for anyone interested in taking up the challenge. Relevant aspects discussed in the article include proposed image segmentation methods, quantitative measures of neuronal morphology, currently available software tools for various related purposes, and morphology databases.

Keywords: Neuron tracing, digital reconstruction, image segmentation, quantitative image analysis, pattern recognition, bioimage informatics, neuroinformatics, neuroscience, software tools, databases.

INTRODUCTION

Understanding how the brain works is undoubtedly one of the grandest challenges for modern science (1). Acquiring profound knowledge of the structure, function, and development of the nervous system at the molecular, cellular, and systems levels is of crucial importance in this endeavor, as processes at these levels are intricately linked with higher-order cognitive functions (2–4) and are the primary targets of drugs and therapies in treating neurological and psychiatric disorders. Since research in the various areas of neuroscience is increasingly relying on imaging, giving rise to massive amounts of heterogeneous and complex data collected at multiple scales of observation, the need for advanced bioimage informatics (5) and neuroinformatics (6–8) solutions for integrating and analyzing these data is rising rapidly.

Of particular importance in this context is the development of computational methods and tools for the study of neuronal anatomy (9). The morphological properties of the axonal and dendritic arborizations constitute a key aspect of neuronal phenotype and play a determinant role in network connectivity (10). Their quantitative analysis enables studying the intrinsic and extrinsic factors influencing neuronal development,
the neuropathological changes related to specific syndromes, the relationships between neuronal shape and function (11), and the effects of specific compounds, providing invaluable information for drug development. This requires converting the (often large and sparse) image data sets acquired for such studies into a much more parsimonious representation of the neuronal topology and geometry consisting of point coordinates, local thickness, and connectivity between points, which captures the essence of the relevant image information and is much easier to archive, exchange, analyze, and compare (12, 13). In addition, these digital reconstructions can be used for simulation of electrophysiological behavior, and for statistical analyses aiming at the development of algorithmic descriptions of neurons.
First attempts to obtain digital reconstructions of neuronal morphology with the help of computers date back at least 45 years (14). These consisted in using the computer to interact with the microscope and to store point coordinates indicated manually by a human operator. In the two decades that followed, many attempts were made to reduce the amount of manual labor involved (15), but the level of automation remained very limited due to a lack of computer power and sophistication in computerized image analysis methods (16). It was recognized to be a difficult problem (17, 18) (Fig. 1), as it involves instructing the computer to mimic some of the complex functions performed by the human visual system. Over the past two decades, the fields of computer science and computer vision have come a long way, with dramatic improvements in both computational power and sophistication. However, while most computer scientists would now deem the problem completely solvable, and commercial as well as academic tools have been developed claiming successful solutions, most neuroanatomists struggle with the lack of general applicability of available tools, and in practice neuronal reconstructions are often still made manually by human experts (10).

The lack of powerful computational tools for automated neuron tracing and reconstruction has inspired several institutes in the past year to establish a competition (with a total of $75,000 in prize money) to encourage the development of new algorithms to advance the field (19). Although it is not to be expected, in view of the past decades of many dedicated efforts, that a one-year competition will lead to the holy grail, it will likely serve to attract new generations of computer scientists to take up the challenge. The purpose of the present article is to support and facilitate this by surveying recent developments in the field and guiding the interested researcher to the relevant literature. Due to space limitations, the overview is necessarily very condensed, but the reader is provided with the information needed to pick up the subject quickly. The subsequent sections discuss image processing methods for segmentation of neuronal structures (specifically the soma, neurites, and spines), proposed measures to quantify these structures, available software (especially freeware) tools for this purpose, and databases to promote and facilitate the exchange of neuromorphological data, thus putting the neuron tracing task into the grander perspective of neuroinformatics.

**SEGMENTATION METHODS**

In contrast with early approaches to neuron tracing using specialized computer controlled microscope systems, which stored only the morphological features measured directly from the imaged samples but not the images themselves, the preferred way nowadays is to first acquire the full image data, as it guarantees a permanent record of the original samples and allows the use of more flexible and more powerful data processing methods (12). The first and by far most critical task in creating digital reconstructions of neurons is image segmentation (the process of assigning to each image element a label indicating to which segment or part it belongs). This usually involves four processing steps (Fig. 2), starting with image preprocessing, and proceeding with the actual segmentation of the cell body (soma), the neuronal tree (axon and dendrites, collectively referred to as “neurites” hereafter), and finally the spines, as briefly discussed next.

**Image Preprocessing**

Neurons may be captured using a variety of specimen preparation and imaging protocols. The choice for a particular protocol is often determined by experimental requirements and practical limitations. However, if accurate reconstruction and quantification is an important goal, attempts should be made to optimize the experimental design from the perspective of subsequently applied image processing and analysis algorithms. Automated image segmentation is hampered by noise (inevitable statistical fluctuations as well as other irrelevant structures), low resolution (ultimately limited by diffraction), inhomogeneous contrast (non-perfect
distribution of the dye), and background gradients (non-uniform illumination). Minimization of these artifacts within the boundary conditions imposed by a given experiment is of key importance (20,21). Although prevention is better than cure, to some extent artifacts can be removed by image processing operations such as (nonlinear) smoothing (21,22), deconvolution (20,23,24), shading correction (21,22), and morphological filtering (25–27). In applications requiring both high magnification and a large field of view, it is also often necessary to make montages (mosaics of images of partially overlapping fields), which calls for accurate image registration and stitching to avoid discontinuities (28–30).
Soma Segmentation

The soma contains the same organelles found in all animal cells and is responsible for maintenance and keeping the neuron functional (1). As the central component of a neuron, where dendritic signals are joined and passed on, its surface area is one of the variables of importance for electrophysiological modeling, and has been found to correlate (moderately but significantly) with total dendritic length (31). Since it constitutes the root of the axonal and dendritic trees, its segmentation can serve as a starting point for segmenting the latter. Especially in images containing multiple (or even many) neurons, first segmenting the somas may help to determine the most likely origin of detected neurites. In the case of fluorescence microscopy, one approach to facilitate detection and segmentation of the somas is to stain them differently: DAPI staining, for instance, will make the nuclei light up in a separate image channel, so that their locations can be more easily determined and can be used as seed points in segmenting out the complete cell bodies in the complementary channel (32). If only one stain is used, the somas may still stand out, if the concentration levels are higher than in the neurites. In phase-contrast microscopy too, the cell bodies often yield more contrast than their processes, in which case they can be segmented simply by means of intensity thresholding (33, 34). If, on the other hand, the somas cannot be distinguished solely based on intensity, they may be discriminated by applying suitable (binary or grayscale) morphological filtering (35–38). For example, repeated “erosion” operations remove all thin and small image structures (the neurites), after which the larger structures (the somas) can be restored by subsequent “dilation” operations (21, 22). In 2D, these standard filters are computationally fast. In order to save computation time in processing 3D images, it has been proposed to detect the somatic areas in 2D projections of the data (in $x$-$y$, $x$-$z$, and $y$-$z$), which define limiting bounding boxes for refined segmentation in 3D (39).

Tree Segmentation

Methods for segmentation of the neuronal tree can be roughly categorized into “global processing” versus “local exploration” approaches (36,39,40). Most global processing algorithms implement the following sequence of operations: binarization, skeletonization, rectification, and graph representation. The binarization step, which aims to yield an initial segmentation of the target image structures, is usually implemented by some form of (adaptive) thresholding (25,26,35,41–49). However, intensity thresholding, while commonly used for its simplicity and efficiency, is generally known to be one of the most error-prone segmentation methods, and it will be successful only if the staining is sufficiently homogeneous, such that the intensity levels in the structures of interest are significantly and consistently different from the background. Alternative approaches, based on contour segmentation, have also been proposed (34, 50–52).

To obtain a more compact description of the neuronal tree, a common next step is to extract the centerlines of the segmented areas (25, 35, 36, 38, 41, 42, 45–48, 51, 53–57), for which various skeletonization algorithms have been proposed. Neurite centerlines may alternatively be obtained directly from the grayscale images, by applying Hessian (58–61) or Jacobian (62) based analysis of critical points, or by nonmaximum suppression (37, 63). The result of the skeletonization step often contains errors (such as spurious gaps or branches) and (especially in 2D) ambiguities (spurious loops or crossings). Various filling and pruning strategies have been developed to try to rectify these retrospectively (25, 37, 38, 46–48, 53, 57, 59, 62, 63). Given the centerlines, it is easy to estimate the neurite diameter at each location, for example using the Rayburst algorithm (46, 64). The final step is to build a graph-theoretic representation by identifying and keeping only the critical points (terminations, bifurcations, inflections) (36, 38, 53).

The second category of tree segmentation methods consists of algorithms which explore an image only locally around relevant structures rather than processing the entire image. It is especially these algorithms to
which the term “tracing” applies. There are at least two reasons for preferring local exploration over global processing algorithms: the latter (due to their global nature) usually work well only in uniformly high-quality (high contrast-to-noise) images, and they are computationally rather wasteful (especially in 3D only a fraction of the image data contains relevant structures). Contrary to global methods, where critical points are usually identified only in the last stages, local methods often start with the detection of topologically relevant points (either manually or using heuristic automatic detection schemes).

The basic idea of local tracing algorithms is to iteratively predict the next point on the neurite (based on information at the current point), and to update this estimate by fitting a (shape or profile) model (39, 65). In high-magnification 3D images of neurites running largely in the axial direction, region-growing methods may be used to segment the neurites in one optical section, whose centroid positions can serve as seed points to initiate segmentation in the next section (66, 67), reminiscent of mean-shift tracking (68, 69) and active-contour based propagation approaches (70–72). More robustness can be expected from algorithms that constrain the search to given start and end points, by defining a cost or “energy” function that assigns a penalty to connecting any two neighboring points (computed from local image features at these points), and minimizing the cumulative cost from start to end point (27, 58, 73–76). A related approach is to fit an active-contour model (based on generalized cylinders) to the image data between given crucial points (77, 78). Such energy minimization approaches are not only suitable for automated tracing, but are also ideal for interactive tracing, as they enable robust linking of image structures while leaving the selection of critical points (bifurcations, crossings, terminations) to the expert user.

**Spine Segmentation**

Although many studies focus on neuronal morphology at the level of the dendritic tree, there is increasing interest in analyzing the structure and function of dendritic spines (79, 80), as they play a major role in the formation and preservation of proper connections between neurons, and thus the ability of the brain to process information. Since the size of these membranous protrusions, in particular their connections to the neuron, are close to the optical resolution limit, deconvolution is often an important preprocessing step (49, 55, 61, 81). But even after deconvolution, spines may appear disconnected (Fig. 1). Algorithms for spine segmentation therefore usually distinguish between attached and (visually) detached spines. Applying heuristic criteria, the former can be detected as small protrusions from the dendritic backbone, obtained from the centerline (skeleton) representation (45, 54, 55, 60, 62, 81, 82), or as peaks in the projected intensity profile of a dendrite (83). Detached spines can be detected as isolated blobs or small segments up to some maximum distance from the backbone (55, 81, 82). Methods for reattaching detached spines to the dendrites have also been proposed (49, 61, 84). Accurate segmentation of the spines after initial detection can be accomplished by applying approaches based on level-sets (49, 61) or Rayburst sampling (46, 64, 84).

**QUANTITATIVE MEASURES**

The advantage of neuronal reconstructions (see examples in Fig. 3) over the raw image data is that they contain all relevant structural information in a condensed form that allows easy computation and statistical analysis of a plethora of quantitative measures of biological variables (13). Even though in practice, due to imaging and image segmentation artifacts, reconstructions may not always be perfectly accurate representations of the true biological data, and the issue of data quality control and addressing “morphological noise” (which can be significant) remains very important (85–88), they constitute our only means of obtaining quantitative results. There exist several ways to categorize quantitative measures of neuronal structure.
Figure 3: Sample neuronal reconstructions selected from the NeuroMorpho.org database based on (principal) cell type: (A) ganglion cell (mouse retina), (B) granule cell (rat hippocampus), (C) motoneuron (cat spinal chord), (D) Purkinje cell (mouse cerebellum), (E) pyramidal cell (rat cerebral cortex), (F) stellate cell (rat cerebral cortex), (G) tangential cell (blowfly visual lobe), (H) relay cell or interneuron (mouse dorsal thalamus). Apart from obvious differences in the arborizations of different cell types, the renderings also reveal a diversity in the reconstructions, which sometimes lack a cell body or axon. The vast majority of the reconstructions in the database were obtained by manual neuron tracing using either the commercial Neurolucida package or the freeware tool NeuronMorpho. All reconstructions in the database are available in the original file format and the standard SWC format. The samples shown here were rendered using the freeware visualization tool V3D (see text and Table 1 for details).

For example, a distinction can be made between topological measures (ignoring metrical dimensions and concerning only the connectivity pattern) versus metrical measures (concerning actual physical “distances”), or between “whole-tree” versus “within-tree” measures (31), or between mathematical types such as differential geometry, symmetry axes, and complexity (11). The subdivision in this section follows the pattern in the previous section, starting with the soma, and subsequently discussing measures to quantify the tree, and the spines. Needless to say, many possible correlations between measures at these different levels could be studied, but these are not explicitly discussed here.

Soma Quantification

An obvious way to quantify a neuronal cell body is by its size, in terms of volume or surface area (either total, or projected, or cross-sectional), and in many studies somatic size is estimated for a first characterization (31). Although size has been shown to correlate with several other neuronal features, its discriminative power is very limited (a given volume can theoretically correspond to infinitely many different shapes). Also, it may be difficult in practice to estimate somatic size accurately, due to ambiguities in the transitions from soma to dendrites (31, 89). In order to characterize somatic shape, important measures include the maximum diameter (major axis), the minimum diameter (minor axis), and especially their ratio. Together
with volume (area in 2D) and surface area (perimeter in 2D), these may be used to compute certain form factors that express how close the shape of a cell body is to representative elementary geometrical shapes. Neurons in the central nervous system show a variety of somatic shapes (also depending on the type or arborization), and relevant quantitative measures to classify these include sphericity (circularity), ellipticity (eccentricity), pyramidalicity (triangularity), rhomboidalicity, and elongation (35).

**Tree Quantification**

Quantitative analysis of the neuronal (in particular dendritic) tree may be performed at different levels of detail. In some studies, for example in evaluating growth factors in explant cultures, where excessive numbers of neurites are observed with many visual ambiguities, it may suffice to measure total neurite count or even just the total area covered by the segmented regions (34,43). Other global measures frequently encountered in the literature include the total height, width, depth, length, and volume of the tree. A still popular method for measuring the spatial distribution and extension of the tree in some more detail is Sholl analysis (90). In this analysis, the number of intersections of neuronal processes with concentric spheres (in 3D) or circles (in the analogous 2D method by Eayrs) (91) of increasing radii centered in the cell soma are counted, and plotted to obtain the so-called Sholl profile of the neuron. Very similar profiles can be obtained more efficiently by counting the (usually considerably smaller) number of branch and terminal points instead (92). Alternative stereological procedures based on counting intersections with other patterns of test lines have also been proposed (93,94). Sholl analysis has been criticized because of its intermingling of topological and metrical parameters, its limited sensitivity (it will detect only relatively large structural differences between groups of neurons), and the fact that it completely ignores orientation differences (31, 89).

Preferential orientation and elongation of the dendritic field can be computed from the direction and the ratio of the principal axes through the soma, or from the angular distribution of the dendritic intersections with the concentric spheres (circles) in Sholl analysis, or from grid density analysis methods (89). Other measures summarizing properties of the whole tree include the total number of (primary, secondary, tertiary) dendrites, their average path length (from terminal tip to origin along the segments), and their radial distances (from terminal tip straight to the somatic origin). Zooming in on the dendritic segments and their bifurcations allows the computation of local metrical measures such as segment length, membrane area, diameter (and ratios thereof comparing the different segments of a bifurcation), bifurcation angles (31) or amplitude (12), taper, contraction, curvature, bending energy, and (multiscale) fractal dimension (11).

From a topological perspective, neuronal trees are essentially binary structures, with each branch point (also referred to as a “node” or “vertex” in graph theory) giving rise to two subtrees. The number of subsequent nodes in both subtrees can be used to compute the so-called partition asymmetry index, whose mean value over all nodes can be taken as a measure of tree asymmetry (31, 95). Various alternative measures of asymmetry have been proposed, in particular “caulescence” (weighted partition asymmetry of nodes along the main path, maximizing a given metric), which provides a clearer functional consequence (96). The power of these measures lies in the fact that, much more than any of the previously mentioned measures, they account for the characteristic connectivity and branching patterns of the neuronal trees. Another recently proposed measure to compute the (dis)similarity between any two neurons, known as the constrained tree-edit-distance (97, 98), takes this idea even a step further. Essentially, it computes the distance between two node-labeled (unordered) trees as the sum of weighted edit operations (label substitutions, node insertions, and deletions) needed to transform (match) one tree exactly into the other, minimized over all feasible edit sequences. The type of (dis)similarity computed by this measure is determined by the definition of the node labels (local geometrical or topological properties) and of the weight function.
Spine Quantification

In spite of the difficulties associated with the segmentation of dendritic spines, due to their small dimensions and heterogeneity, quantitative analysis of these structures remains important for studies related to neuronal computation (31). Spines of different classes of neuronal cells may show differences in size, shape, distribution, development, complement of organelles, and the receptors they bear (80). Notwithstanding their great morphological diversity, spines are traditionally divided into three gross categories, based on the relative sizes of the (bulbous) spine head and the (usually narrower) neck (80): “mushroom” spines (having a relatively large head and a narrow neck), “thin” spines (having a relatively small head and a narrow neck), and “stubby” spines (having no clear constriction between the head and the attachment to the dendritic shaft). However, since spines show a continuum of shape variations over this classification, and no standardized quantitative criteria exist (yet) to differentiate them, the boundaries between these categories are somewhat arbitrary and may differ between studies (55). Measures for spine quantification and classification encountered in the literature include (on the local level) length, head and neck diameters, orientation, moments, volume, and (on the dendrite level) position, count, and density (45, 49, 55, 60, 81, 84).

SOFTWARE TOOLS

There is an increasing tendency in neuroinformatics (99), and in various other fields of science and engineering (100), to not only publish new ideas but also the software tools developed to test these ideas. Dozens of software tools can be found in the recent literature for performing various tasks in neuroinformatics. In this section, well-known tools (in particular freeware) are surveyed and categorized based on their primary functionality: segmentation, visualization, editing, quantification, generation, simulation, and format conversion. A quick reference of the freeware tools discussed in this section (and more), with corresponding literature references and web links containing more detailed information, is given in Table 1. As indicated in the table, in addition to precompiled binaries for common computing platforms, the source code is also available for an increasing number of these tools, under different open-source licenses.

Segmentation Tools

One of the earliest software tools for 3D neuron tracing is the widespread (commercial) Neurolucida package (MBF Bioscience, USA), which originally supported only manual operation (101), but was later extended with the AutoNeuron module supporting automatic tracing. Other prominent commercial software tools providing advanced manual and (semi)automatic tracing functionality include the FilamentTracer (formerly NeuronTracer) module of the Imaris package (Bitplane, Switzerland) and (especially for 2D high-content screening applications) HCA-Vision (CSIRO Biotech Imaging, Australia) (37). Nowadays, many commercial software packages for (live) cell imaging also contain modules for neuron segmentation and analysis (not mentioned explicitly here), with varying levels of sophistication.

Quite a number of freeware tools exist that offer similar functionality. An example of a freeware tool for manual 3D neuron tracing is the Neuron_Morpho plugin (12) to ImageJ (NIH, USA), an open-source Java-based image processing and analysis platform (102, 103). Semiautomatic tracing of neurites (or similar structures) through optical sections may also be done using the Reconstruct tool (66, 67). For measuring neurite length parameters in 2D assays, a popular semiautomatic tool is the NeuronJ plugin to ImageJ, whose corresponding paper (58) was the highest cited paper in the past decade of all those discussed in the present survey (according to the ISI Web of Science). NeuronJ has been used as a reference tool for testing later 2D neuron tracing algorithms (32, 37, 48) and has inspired others to develop interactive tools such as
<table>
<thead>
<tr>
<th>TOOL</th>
<th>TASKS</th>
<th>SOURCE</th>
<th>PLATFORM</th>
<th>LINK</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cvapp</td>
<td>V E</td>
<td>C</td>
<td>Yes</td>
<td>Java <a href="http://www.compneuro.org/CDROM/docs/cvapp.html">http://www.compneuro.org/CDROM/docs/cvapp.html</a></td>
</tr>
<tr>
<td>Eutectic (18)</td>
<td>S V E</td>
<td>Q</td>
<td>Win</td>
<td><a href="http://bellsouthpwp.net/c/a/capowski/NTSPublic.html">http://bellsouthpwp.net/c/a/capowski/NTSPublic.html</a></td>
</tr>
<tr>
<td>FARSIGHT (105)</td>
<td>S V E</td>
<td>Q</td>
<td>Win/Mac</td>
<td><a href="http://www.farsight-toolkit.org/">http://www.farsight-toolkit.org/</a></td>
</tr>
<tr>
<td>GENESIS (117)</td>
<td>S V E</td>
<td>Q</td>
<td>Win/Lin/Mac</td>
<td><a href="http://www.genesis-sim.org/">http://www.genesis-sim.org/</a></td>
</tr>
<tr>
<td>L-Neuron (110)</td>
<td>G</td>
<td>Yes</td>
<td>Win/Lin</td>
<td><a href="http://krasnow.gmu.edu/cn3/L-Neuron/">http://krasnow.gmu.edu/cn3/L-Neuron/</a></td>
</tr>
<tr>
<td>NETMORPH (115)</td>
<td>G</td>
<td>Yes</td>
<td>Source</td>
<td><a href="http://www.netmorph.org/">http://www.netmorph.org/</a></td>
</tr>
<tr>
<td>NeuGen (113)</td>
<td>V G</td>
<td></td>
<td>Win/Lin/Mac</td>
<td><a href="http://www.neugen.org/">http://www.neugen.org/</a></td>
</tr>
<tr>
<td>NeuRA</td>
<td>S</td>
<td>Yes</td>
<td>Source</td>
<td><a href="http://www.neura.org/">http://www.neura.org/</a></td>
</tr>
<tr>
<td>neuroConstruct (114)</td>
<td>V E G</td>
<td>Yes</td>
<td>Java</td>
<td><a href="http://www.neuroconstruct.org/">http://www.neuroconstruct.org/</a></td>
</tr>
<tr>
<td>NeuroGL</td>
<td>V</td>
<td></td>
<td>Win</td>
<td><a href="http://www.mssm.edu/cnic/tools-neurogl.html">http://www.mssm.edu/cnic/tools-neurogl.html</a></td>
</tr>
<tr>
<td>Neuromantic (75)</td>
<td>S V E</td>
<td>Q</td>
<td>Win</td>
<td><a href="http://www.reading.ac.uk/neuromantic/">http://www.reading.ac.uk/neuromantic/</a></td>
</tr>
<tr>
<td>NeuroML Validator</td>
<td></td>
<td></td>
<td>C</td>
<td>Browser <a href="http://www.neuroml.org/">http://www.neuroml.org/</a></td>
</tr>
<tr>
<td>NEURON (118)</td>
<td></td>
<td></td>
<td>M</td>
<td>Yes Win/Lin/Mac <a href="http://www.neuron.yale.edu/neuron/">http://www.neuron.yale.edu/neuron/</a></td>
</tr>
<tr>
<td>Neuron_Morpho (12)</td>
<td>S E</td>
<td>Yes</td>
<td>ImageJ</td>
<td><a href="http://www.personal.soton.ac.uk/dales/morpho/">http://www.personal.soton.ac.uk/dales/morpho/</a></td>
</tr>
<tr>
<td>NeuronCyto (32)</td>
<td>S Q</td>
<td>Yes</td>
<td>Matlab</td>
<td><a href="http://neuroncyto.bii.a-star.edu.sg/">http://neuroncyto.bii.a-star.edu.sg/</a></td>
</tr>
<tr>
<td>NeuronIQ (104)</td>
<td>S V E</td>
<td>Q</td>
<td>Matlab</td>
<td><a href="http://www.cbi-tmhs.org/NeuronIQ/">http://www.cbi-tmhs.org/NeuronIQ/</a></td>
</tr>
<tr>
<td>NLMorphologyConverter</td>
<td></td>
<td></td>
<td>C</td>
<td>Win/Mac <a href="http://www.neuronland.org/">http://www.neuronland.org/</a></td>
</tr>
<tr>
<td>NLMorphologyViewer</td>
<td>V</td>
<td></td>
<td>Win</td>
<td><a href="http://www.neuronland.org/">http://www.neuronland.org/</a></td>
</tr>
<tr>
<td>ORION (76)</td>
<td>S V Q</td>
<td>Win</td>
<td>Win <a href="http://cbl.uh.edu/ORION/research/software/">http://cbl.uh.edu/ORION/research/software/</a></td>
<td></td>
</tr>
<tr>
<td>Reconstruct (66)</td>
<td>S V E</td>
<td>Q</td>
<td>Win</td>
<td><a href="http://tech.groups.yahoo.com/group/reconstruct_users/">http://tech.groups.yahoo.com/group/reconstruct_users/</a></td>
</tr>
<tr>
<td>V3D (106)</td>
<td>S V Q</td>
<td>Win/Lin/Mac</td>
<td><a href="http://penglab.janelia.org/proj/v3d/">http://penglab.janelia.org/proj/v3d/</a></td>
<td></td>
</tr>
<tr>
<td>XL_Calculations (108)</td>
<td>Q</td>
<td>Yes</td>
<td>Java</td>
<td>Available from the developers on request</td>
</tr>
</tbody>
</table>

**Table 1:** Freeware tools for the various tasks in neuroinformatics discussed in the paper. The columns list (from left to right) the tool name (with primary literature reference), main functionalities (the shaded cell indicates the primary task), whether source code is available, the platform(s) on which precompiled versions run, and the web link. Tasks: S = segmentation, V = visualization, E = editing, Q = quantification, G = generation, M = simulation, C = format conversion. Platforms: ImageJ = plugin to ImageJ (runs on all platforms with ImageJ installed), Java = stand-alone Java application (runs on all platforms with the Java Virtual Machine installed), Matlab = application written in the Matlab environment (runs on all platforms with Matlab installed), Win = native Microsoft Windows application, Lin = native Linux application, Mac = native Apple Macintosh application, Source = available only as source files (to be compiled to the platform of choice), Browser = online application (runs on all platforms in a suitable web browser). See the main text and the given literature references and web sites for more detailed information.
Neuromantic (75) for neuron reconstruction in 3D. Higher levels of automation for 2D applications are found in tools such as NeuronMetrics (47), NeuriteTracer (48), and NeuronCyto (32), and for 3D applications in NeuronStudio (46), NeuronIQ (104), ORION (76), and FARSIGHT (105). While most of the tools focus on tree segmentation, several of them, notably Neurolucida, FilamentTracer, FARSIGHT, NeuronIQ, and NeuronStudio, also explicitly detect the somas and/or the spines.

**Visualization Tools**

Although neuron tracing tools usually also enable visualization of the tracing results, there exist various tools developed specifically for this purpose. Most of them accept multiple morphology file formats (see below for more on this issue). An example is Cvapp, which was originally written for inspection and curation of neuronal morphology data, but was later extended to perform a variety of other functions, including conversion between file formats. Other examples of tools for visualization of neuronal morphology data from different formats include NeuroGL and NLMorphologyViewer. A more versatile visualization tool is V3D (106), which supports up to 5D rendering (spatially 3D over time and in multiple colors) of image data as well as surface data, relational data that can be described as a graph (such as neuronal data), point clouds, and landmarks. In addition, it offers various image analysis functions (as add-on modules), and also supports user-developed plugins with which a user can exploit the V3D platform in developing new functions. A similar commercial tool is Amira (Visage Imaging, USA).

**Editing Tools**

Even though full automation of data processing and analysis remains the ultimate goal, and is in fact a prerequisite for high-content/high-throughput experiments, current state of the art algorithms will achieve this only under highly optimal conditions that are hardly ever met in practice. In most cases, “raw” neuron tracing results contain a variety of errors, ranging from gross failures such as incorrectly included or excluded image structures (false positives or negatives), to more subtle flaws such as incorrectly broken or merged structures (false fragmentations or continuations). To what extent these are harmful to subsequent analyses depends greatly on the parameters studied. For example, a false negative segmentation of part of the dendritic tree may greatly affect total counts (of dendrites and spines), but may have negligible effect on averages (such as spine density). Conversely, a false fragmentation of the tree may affect averages (such as dendrite length), but may be relatively harmless to summed quantities (such as total dendrite length). In any experiment, the possibility of tracing errors and their potential effects on subsequent quantitative analyses require careful consideration, and most automated segmentation tools (as indicated in Table 1) therefore support manual postediting of the results to correct for these.

**Quantification Tools**

Similar to visualization and manual postediting of results, quantitative analysis too is supported by most neuron tracing tools (albeit with greatly varying levels of detail and sophistication), but there exist a few tools developed specifically for this task. The most extensive tool in this category is L-Measure (107), which can compute over 100 independent morphological parameters (regarding the soma and the tree), from populations of cells, to individual neurons, to portions thereof. This enables detailed comparative analyses on large numbers of neurons, the discovery of characteristic morphological features across cell classes, the detection of differences induced by specific growth factors, the analysis of developmental changes, the extraction of parameter distributions required for computational simulations to generate virtual neurons, and the assessment of the quality and limitations of these models by comparing their emergent properties with the original experimental data (107). Another tool is XL_Calculations (108), which was designed to
distinguish between neurons at different stages of differentiation, and to this end facilitates batch processing of large numbers of NeuronJ tracing files and the automated computation of over 45 different quantitative tree measures. Quantitative analysis of spine properties is offered by most of the previously mentioned segmentation tools that explicitly detect these structures.

**Generation Tools**

A very challenging goal of neuroinformatics is the development of algorithmic descriptions of neuronal morphologies (109). Not only is this useful for obtaining even more compact representations of neurons than digital reconstructions, it also improves our understanding of the underlying fundamental processes and parameters of nature, and enables automatic generation of unlimited numbers of virtual neurons for modeling and simulation experiments. Several tools already exist for the computer generation of single neuron morphology or even complete neuronal networks based on statistical models derived from large databases of previously extracted true morphologies. Examples include L-Neuron (110, 111) (running local algorithms independent of possible extrinsic constraints) and ArborVitae (111, 112) (not yet available but implementing global algorithms in which populations of neurons grow based on environmental constraints). Other tools for the generation of networks of neurons closely matching the morphology and connectivity of different brain regions include NeuGen (113), neuroConstruct (114), and NETMORPH (115).

**Simulation Tools**

Modeling and simulation of neuronal networks is important for testing hypotheses derived from computational paradigms (116). GENESIS (117) (an acronym of GEneral NEural SImulation System) was one of the first systems developed for the simulation of biologically realistic models at different levels, ranging from subcellular components and biochemical reactions, to single neurons, large networks, and complete systems models. An analogous simulation environment is NEURON (118, 119). Both provide ways to construct, test, and manage models, such that expertise in numerical methods or programming is not required for productive use. A much more complete discussion of these, and six other simulation systems and underlying strategies can be found in (120), while the issue of interoperability between simulators is discussed in (121). Similar to quantification of morphological properties, the simulation of electrotonic and electrophysiological behavior too may be affected significantly by variabilities in the digital reconstruction of neurons, emphasizing the importance of accurate and consistent reconstruction (86, 88).

**Conversion Tools**

In view of the different specialized tasks in neuroinformatics, neuronal reconstructions should ideally be saved in an open, well-documented, and extensible format, allowing easy archiving and exchange of information, not only between software components but also between different labs to facilitate collaboration. Unfortunately, many (especially commercial) software tools use their own (proprietary) formats. A common open file format is SWC (112, 122), and several tools can convert other formats to this format (L-Measure), or convert it to formats used by simulation tools (Cvapp). However, neuroanatomy is but one type of model information about neural systems. More inclusive model descriptions are possible using NeuroML (123, 124), a neural markup language based on the extensible markup language (XML). MorphML (125) is the first level of the NeuroML standard and deals with neuroanatomical information (higher levels deal with biophysical properties of cells and network connectivity). The most complete tool to date for format conversion (including from/to the MorphML standard) is NLMorphologyConverter.
MORPHOLOGY DATABASES

An important aspect of scientific research in general is the publication of ideas and results, not only to allow reproduction and verification, but also to help increase the rate of scientific progress. Articles in scientific journals often do not serve this goal, but are merely advertisements of new findings (100), still requiring many days (if not weeks or months) of duplicate work for others to be able to use them productively. The previous section already discussed the issue of software tools to facilitate the reuse of computational methods. Another aspect, briefly discussed in the present section, is the sharing of the actual experimental data underlying a scientific publication. The rationale here is that the potential of the whole body of shared data is larger than the sum of its parts. In this context, a distinction must be made between primary data (usually the raw images), and secondary data (any information derived from the primary data). For obvious reasons, sharing of the former has been the subject of considerable debate in neuroscience (13, 126–129). On the other hand, the publication of secondary data such as neuronal reconstructions is now gaining support, and first reviews of data sharing in this area have shown rewards in comparative morphometric analysis, statistical estimation of potential synaptic connectivity, anatomically accurate electrophysiological simulation, and computational modeling of neuronal shape and development (129).

Data sharing is best implemented in the form of publically accessible databases. In the case of digital reconstructions of neuronal morphology, the most comprehensive database is NeuroMorpho.org (13, 130), an online (its name derives straight from its URL) and centrally curated inventory containing contributions from dozens of labs. The database can be browsed by animal species, brain region, cell type (see Fig. 3), or contributing lab, and is also searchable by general keywords, or more specifically by morphometry or metadata, including information on the imaging protocol and parameters as well as the method used for reconstruction. Being the largest web-accessible collection of neuronal reconstructions and associated metadata, it is one of the integrated resources in the Neuroscience Information Framework (NIF), a recent initiative sponsored by the National Institutes of Health (NIH) for integrating access to and promoting the use of web-based neuroscience resources (131, 132). Moreover, it mirrors many other repositories, in line with its primary goal of achieving and maintaining dense coverage of all publically available neuronal reconstructions. The increasing online availability of digital data through such initiatives will create unprecedented opportunities for data mining and will likely accelerate novel scientific discoveries.

CONCLUSIONS

As in many areas of the health and life sciences, the processing, analysis, and management of images, metadata, and results from neuroscientific experiments relies increasingly on automated methods and tools. From the perspective of neuroinformatics at large, a key task is the reconstruction and study of neuronal morphology from image data, and we have surveyed the essential methods (for segmentation and quantification) and tools (software packages and public databases) developed for this purpose. Notwithstanding many praiseworthy efforts in the past four decades, the field has only just begun to show its potential, and challenges remain in all relevant aspects, as briefly summarized in these concluding remarks.

For image segmentation (the most critical step in obtaining digital reconstructions), many approaches have been proposed, but a general purpose method of sufficient robustness to deal with the large variability of image data from different labs and experiments is yet to be developed. The DIADEM challenge (19) will prove important in this regard, as it stimulates not only the development of new algorithms, but also their careful validation and performance assessment, by offering representative data sets and expert manual re-
constructions. From general observations in related areas of biomedical image analysis (133) it would seem that improved methods for neuron tracing are possible by bridging the dichotomy of global processing versus local exploration algorithms, exploiting the best of both worlds in a multiscale framework, and moving towards probabilistic rather than deterministic image segmentation approaches.

Another very important aspect of developing improved image segmentation algorithms is the careful design of parameters. A translation of Occam’s razor principle to this problem suggests that ending up with a large number of user-settable parameters is indicative of poor algorithm design. In order to handle variability, parameters are hard to avoid in general, but the least algorithm designers should strive to accomplish is to make them (bio)physically meaningful and independent of underlying technical issues. The next level of sophistication, which deserves considerably more attention in the future, is the automation of parameter selection. First efforts in this direction (134) have already shown promising results.

For the quantification of neuronal morphology based on the reconstructions, many sensible measures have been proposed, but most of the ones commonly studied in practice are rather rudimentary, and the question remains to what extent the present spectrum of measures really captures the essential (spatial) features and their relation to (temporal) function. In this context too, Occam’s razor principle deserves consideration. Scientific inquiries can be expected to become more complex and more inclusive, and the integration of information from multiple levels of observation will call for quantitative measures reflecting this.

Finally, in terms of software development, apart from the implementation of more robust segmentation and quantification algorithms, a major challenge lies in improving the integration and interoperability of different components. Ideally, this should result in a platform enabling researchers to perform the full range of data collection, processing, and analysis tasks in a given experiment, without switching tools or converting data. Together with growing databases of neuronal reconstructions (13), these solutions will in the near future enable effective and efficient investigations of much larger scale and scope.

ACKNOWLEDGMENTS

The author gratefully acknowledges the constructive comments of the anonymous reviewers which helped improve the survey. This work was financially supported by the Netherlands Organization for Scientific Research (NWO) through VIDI-Grant 639022401 awarded to the author.

REFERENCES


