Fast Tracing of Microtubule Centerlines in Electron Tomograms

Britta Weber* Zuse Institute Berlin
Marit Möller† Zuse Institute Berlin
Jean-Marc Verbavatz‡ MPI-CBG Dresden
Daniel Baum§ Zuse Institute Berlin
Hans-Christian Hege¶ Zuse Institute Berlin
Steffen Prohaska‖ Zuse Institute Berlin

Index Terms: J.3 [Life and Medical Sciences]: Biology and Genetics I.4.6 [Computing Methodologies]: Image Processing and Computer Vision—Segmentation; I.3.8 [Computing Methodologies]: Computer Graphics—Applications;

1 Introduction

The organization and role of the cytoskeleton, particularly the network formed by *microtubules*, is an active topic of research in cell biology. To study the role of these tubular shaped macromolecules, $25\,nm$ in diameter, information about their length, number, spatial distribution and dynamics must be gathered. Currently, the only technique that provides the needed resolution is electron microscopy. To gather 3d spatial information, samples are cut in sections of $300\,nm$ thickness to allow a projection with an electron beam. The volume is then reconstructed from projections with different tilt angles, a technique called electron tomography. Unfortunately, samples cannot be tilted up to 90° in the microscope, which results in an artifact in electron tomograms known as 'missing wedge'. This artifact as well as angular sampling, imperfections in the sample preparation and destructive effects caused by the electron beam, result in a low signal-to-noise ratio in electron tomograms making their automatic processing particularly challenging. Fig. 1 shows examples of microtubules in electron tomograms.

To draw reliable conclusions about the structure of the microtubule network, e.g. during cell division, microtubules need to be identified and their end types classified precisely in a large number of specimens. Manual analysis is often unfeasible: The midsection of a centrosome in the mitotic spindle of C. Elegans, for example, contains between 600 and 1000 microtubules. With current methods, manual segmentation of centerlines in one of these sections takes 10 to 30 hours depending on the image quality. Therefore, biological research on microtubules using electron tomography is currently often based on single observations as opposed to deriving statistics from many samples. On the other hand, automatic extraction and classification on the centerlines remains an open problem and has not yet been solved with the needed quality (see [2]).

The major goal for any attempt to solve this problem is to reduce the time spent manually without compromising quality. For large scale analyses at least one order of magnitude of speedup is needed. Segmentation and classification can be divided into two tasks: The tracing of the centerlines and the classification of microtubule ends as being open or closed - not a trivial assignment as can be seen in Fig. 1. Since the major bottleneck is the tracing of the centerlines, we describe how the time spent manually on this task can already be reduced significantly by a combination of automatic tracing and manual interaction.

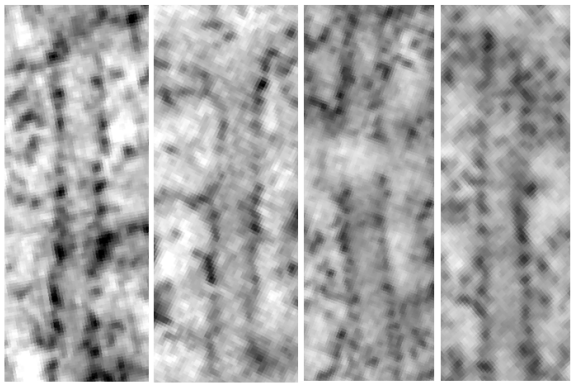

Figure 1: Examples for slices showing lengthwise cuts through microtubules in electron tomograms. From left: Closed end, open end, defect in sample, unclear end.

2 Semi-automatic tracing

The most popular tool for tracing microtubules manually is IMOD [1], which comes with an editor to add lines in tomograms. Navigation in a tomogram is achieved by displaying a slice through the tomogram. Microtubules can be traced by placing points along their path in the tomogram. IMOD does not offer automatic or semi-automatic tracing, but only functionality to manually trace microtubules. For this task, IMOD provides very good usability.

To accelerate tracing, as a first attempt, we developed a semi-automatic tracing tool which automatically traces a single microtubule for a manually determined seed point. For the tracing of a single microtubule, we utilized an algorithm similar to the one described in Rigort et al. [3]. This algorithm combines template matching with a tracing algorithm using the information from template matching. The editor was implemented as an add-on to Amira [4]. We employed an iterative software development process in close collaboration with biologist, gathering and implementing feature requests from experts on a regular basis until no more significant changes were expected.

To estimate the quality of the implemented editor, we compared the performance of experts when using Amira and IMOD in three experiments and measured the time needed for each task: (a) pure tracing of centerline of microtubules without classifying their endpoints for 1 hour, (b) tracing and classifying microtubules for 1 hour, and (c) searching and tracing missing microtubules in a nearly completely segmented network, spending as much time as needed.

The experiments showed that the speed-up when using Amira is not greater than a factor of two. In more detail, experiment (a) showed that the time needed for tracing a microtubule in IMOD had been overestimated: The gained speed-up with our semi-automatic tool was a factor of two. Experiment (b) showed that the classification of endpoints takes as much time as manual tracing of cen-

*e-mail: britta.weber@zib.de
†e-mail: moeller@zib.de
‡e-mail: verbavatz@mpi-cbg.de
§e-mail: baum@zib.de
¶e-mail: hege@zib.de
‖e-mail: prohaska@zib.de

terlines. Experiment (c) showed that searching for missing microtubules has a very high impact: It took experts three times longer to trace microtubules than in experiment (a).

3 Automatic Tracing

Since searching for microtubules seemed to have the highest impact on the timings, in our second attempt, we focused on a fully automatic approach to extract microtubules. Though the literature on the extraction of tubular structures is quite large, according to our experiences the only serious approach to automatically trace microtubules is Nurgaliev et al. [2]. However, they state that their approach, though promising, needs improvement before applying it in biological experiments.

In Rigort et al. [3] we showed that template matching combined with a simple tracing algorithm can be used to automatically trace actin filaments in ice-embedded electron tomograms. We adapted this method to microtubule tracing in plastic-embedded electron tomograms. However, for reliable statistics on the number and lengths of microtubules in a specimen, the quality expectations are higher than for the purposes in reference [3]. Since the outcome of the tracing depends on several parameters, we investigated how the quality of the result differs w.r.t. the parameter settings and variations of image quality that typically occur in electron tomograms (see Torsney-Weir et al. [5] for a description of the analysis and appropriate parameter settings).

To measure the quality of the automatic tracing, we point-wisely compared our results to manual tracings using two measures, Precision and Recall. Here, a point on a line of the automatic tracing is considered true positive (TP) if a point on a line in the manual tracing within a radius of $10\,nm$ (about the radius of a microtubule) can be found. Else, this point is counted as false positive (FP). Unmatched points in the manual tracing are considered false negatives (FN). Precision is then computed as $\#TP/(\#TP+\#FP)$ and Recall as $\#TP/(\#TP+\#FN)$. For our comparisons, the line sets were sampled at a distance of one nanometer. Using Precision and Recall, we searched for stable parameter settings of the algorithm (see [5]) by comparing automated to manual tracings on five datasets where a ground truth was generated from four manual tracings for each. The chosen parameter settings were then evaluated on 27 tomograms for which manual tracings were available (see Fig. 2 for an example). On average, Precision was 0.88 and Recall 0.94. The 12% false positives can partly be explained by incomplete manual tracings.

The results indicate that the automatic approach might be useful for answering simple questions about the amount of microtubules in a specimen. However, a closer inspection of the resulting lines revealed that, although only few parts of lines were false, the automatic approach often created broken lines and very short false positives. Such defects cannot be detected by the above described point-wise comparison. Lengths and numbers derived from the automatic tracing are therefore unreliable.

4 Combining Both Approaches

Though imperfect, the automatic approach has the advantage that no time at all is spend manually once the best parameter configuration is found. However, for a detailed analysis of length or number in a microtubule network, the result must be corrected.

The natural question arising is whether a manual correction of the automatic results is feasible. A first test on few tomograms showed that it takes approximately 1.5 hours to correct an automatic tracing resulting in 830 lines after correction. In experiment (a), users traced between 388 lines on average in one hour, a number that would decrease the more lines are found as indicated by results of experiment (c). We therefore assume that correcting an automatic tracing is 3 to 4 times faster than tracing only manually with conventional methods. However, users refused to correct automatic tracings in the presence of many false positives or many false negatives, saying it is too annoying correcting something that is utterly false. The correct parameter choice is therefore important to gain this factor.

5 Conclusion

Although the hoped-for order of magnitude increase in speed has not been reached yet, we believe that the next problem to be solved is the classification of microtubule endpoints. Since we measured that this step takes as long as tracing microtubules, a further improvement of the tracing of microtubule centerlines will not result in the acceleration needed. We believe that this problem is even more challenging than tracing of centerlines (see Fig. 1). To our knowledge, no automatic method exists for reliable endpoint classification of microtubules.

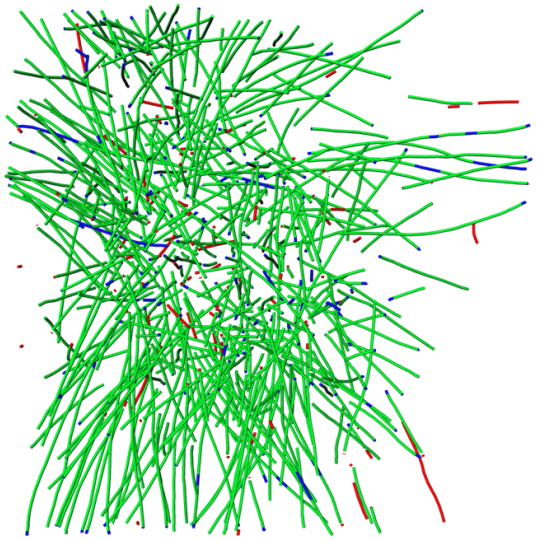

Figure 2: Visual comparison of manual and automatic tracing. Green: True positives. Red: False positives. Blue: False negatives.

Acknowledgements

This project was funded by the Max Planck Society. We thank Annett Boden, Gunnar Fabig and Eric Seemann of the MPI-CBG Dresden, whose tireless testing and patient writing of bug reports made this work possible. We also thank Garrett Greenan and Thomas Müller-Reichert for providing the data used in the experiments. Finally, we thank Tony Hyman for his constant support on this project.

References

[1] J. R. Kremer, D. N. Mastronarde, and J. R. McIntosh. Computer visualization of three-dimensional image data using IMOD. *Journal of Structural Biology*, 116(1):71–76, January 1996.

[2] D. Nurgaliev, T. Gatanov, and D. J. Needleman. Automated identification of microtubules in cellular electron tomography. *Methods Cell Biol*, 97:475–95, 2010.

[3] A. Rigort, D. Günther, R. Hegerl, D. Baum, B. Weber, S. Prohaska, O. Medalia, W. Baumeister, and H.-C. Hege. Automated segmentation of electron tomograms for a quantitative description of actin filament networks. *submitted to Journal of Structural Biology*, June 2011.

[4] D. Stalling, M. Westerhoff, and H.-C. Hege. Amira: A highly interactive system for visual data analysis. In C. D. Hansen and C. R. Johnson, editors, *The Visualization Handbook*. Elsevier, 2005.

[5] T. Torsney-Weir, A. Saad, T. Möller, H.-C. Hege, B. Weber, and J.-M. Verbavatz. Tuner: Principled parameter finding for image segmentation algorithms using visual response surface exploration. *accepted at IEEE Conference on Visualization (VIS-2011)*.