Journal of Structural Biology xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Journal of Structural Biology

journal homepage: www.elsevier.com/locate/yjsbi

Automated tilt series alignment and tomographic reconstruction in IMOD

David N. Mastronarde *, Susannah R. Held

Department of Molecular, Cellular, and Developmental Biology, University of Colorado, Boulder, CO 80309, United States

ARTICLE INFO

Article history:
Received 12 May 2016
Received in revised form 15 July 2016
Accepted 18 July 2016
Available online xxxx

Keywords:
Electron tomography
Tilt series alignment
Tomographic reconstruction
Automated processing

ABSTRACT

Automated tomographic reconstruction is now possible in the IMOD software package, including the merging of tomograms taken around two orthogonal axes. Several developments enable the production of high-quality tomograms. When using fiducial markers for alignment, the markers to be tracked through the series are chosen automatically; if there is an excess of markers available, a well-distributed subset is selected that is most likely to track well. Marker positions are refined by applying an edge-enhancing Sobel filter, which results in a 20% improvement in alignment error for plastic-embedded samples and 10% for frozen-hydrated samples. Robust fitting, in which outlying points are given less or no weight in computing the fitting error, is used to obtain an alignment solution, so that aberrant points from the automated tracking can have little effect on the alignment. When merging two dual-axis tomograms, the alignment between them is refined from correlations between local patches; a measure of structure was developed so that patches with insufficient structure to give accurate correlations can now be excluded automatically. We have also developed a script for running all steps in the reconstruction process with a flexible mechanism for setting parameters, and we have added a user interface for batch processing of tilt series to the Etomo program in IMOD. Batch processing is fully compatible with interactive processing and can increase efficiency even when the automation is not fully successful, because users can focus their effort on the steps that require manual intervention.

1. Introduction

Electron tomography (ET) has historically been time-consuming, both for acquisition and for alignment of images and reconstruction of a tomogram. Automating these processes and improving overall throughput has been a major goal since it became possible to acquire images with digital cameras (Fung et al., 1996; Koster et al., 1992). Improvements in microscope performance and development of automated acquisition (Mastronarde, 2005; Suloway et al., 2009; Zheng et al., 2004) have made it possible to acquire many gigabytes per day of tilt series data and thus to pursue research that requires large amounts of data. Tomography on frozen-hydrated samples (cryoET) is often destined for sub-volume averaging (for recent reviews, see (Asano et al., 2016; Briggs, 2013)), which has recently reached sub-nanometer resolution (Bharat et al., 2015; Pfeffer et al., 2015; Schur et al., 2013). Although various factors besides the number of particles limit the resolution of such averaging, such as the quality of the tilt series alignment, the efficiency of the camera in capturing high-resolution information, and the complexity and flexibility of the structure being averaged, having many particles from many tomograms was a key ingredient in the first of these studies (Schur et al., 2013). High-throughput tomography on plastic-embedded samples allows multiple strains or experimental conditions to be compared with a statistically meaningful number of samples (e.g., Nannas et al. (2014)); it also enables ambitious studies of very large volumes (~50 μm^3 at 1.5 nm voxel, Hoog et al. (2007), ~200 μm^3 at ~5 nm voxel, Noske et al. (2008), ~100 μm^3 at 1.2 nm voxel, Redemann and Muller-Reichert (2013)). For either kind of sample, high-throughput tomography at the cellular level enables informative comparative studies (Ding et al., 2015).

The IMOD software package (Kremer et al., 1996; Mastronarde, 1997) contains a comprehensive set of programs for tomographic

Abbreviations: CryoET, cryo-electron tomography; ET, electron tomography; MADN, normalized median absolute deviation from median; SD, standard deviation; SNR, signal-to-noise ratio; CCC, cross-correlation coefficient; 3-D, three-dimensional.

* Corresponding author at: Dept. of MCD Biology, University of Colorado, 347 UCB, Boulder, CO 80309, United States.

E-mail address: mast@colorado.edu (D.N. Mastronarde).

http://dx.doi.org/10.1016/j.jsb.2016.07.011
1047-8477/

reconstruction and a graphical user interface (Etomo) that manages the whole process, from removal of artifacts in the raw images to trimming and scaling of the final volume. It was developed in a national center for 3-D electron microscopy where tomography was done on a relatively diverse range of plastic-embedded and frozen-hydrated specimens. It has been widely used elsewhere with an even broader range of specimens and acquisition protocols. Problems have arisen in particular data sets, and the software has been improved either to handle more difficult sets as a matter of course, or to provide a special parameter or tool by which the user can resolve a problem. For any step that runs automatically, there is a fallback procedure so the user can get through the step if the automation fails. IMOD's user interface provides access to those settings, tools, and procedures; although they are often hidden during ordinary use, they are available in its advanced display mode. The general approach has been to provide some way to rescue almost every data set. Although many cryoET projects can tolerate considerably less than 100% yield of usable tomograms, in some projects with plastic-embedded specimens there is a strong motivation to get a usable result from every data set, such as when serial sections are being reconstructed.

Every step in IMOD's reconstruction process can now be run fully automatically to obtain a high-quality tomogram. This automation has been implemented within the existing processing framework and in a way that is compatible with interactive processing through Etomo. If there is a failure, users can resolve the problem interactively then resume automated processing. Thus, these capabilities can improve the efficiency of dealing with every tilt series, even the ones that require intervention. The flexibility of the user interface for batch processing makes computing tomographic reconstructions truly high-throughput.

This paper describes the key recent developments that have enabled this automation: (1) improvements in obtaining a model of fiducial marker positions and in solving for the tilt series alignment from those positions; (2) application of methods for determining the location of structure in a tomogram to several problems, including the accurate alignment of the two tomograms from dual-axis tilt series (Mastronarde, 1997; Penczek et al., 1995); and (3) the creation of tools that manage the automated processing. Because this paper covers a wide range of topics, there is an integrated presentation of motivation, methods, and relevant results for each topic in turn. Some of this work has been reported in abstract form (Mastronarde, 2013).

The software is freely available at http://bio3d.colorado.edu/imod. The improvements in fiducial alignment and a program to manage batch processing were available in the IMOD 4.7 stable release version. Beta releases since then have included significant refinements in the latter program plus the other features described here; for any batch processing, version 4.8.51 or higher should be used. Further details on the operations performed by programs described here can be found in the manual pages, accessible from http://bio3d.colorado.edu/imod/betaDoc/program_listing.html.

2. Methods and results

2.1. Automatic generation of a seed model

The approach in IMOD for constructing a fiducial model has been to start with a "seed model" consisting of selected gold beads near 0° tilt and track those beads from one projection to the next with the program Beadtrack. An advantage of this method is that a faint or obscured bead can be detected using the a priori knowledge of where it should be, even if it would not meet some threshold for detection when attempting to find all beads on an image. IMOD already contained a program for automatically constructing the fiducial model, version 3.0 of RAPTOR (Amat et al., 2008), and some laboratories have relied on it extensively, particularly for cryoET. However, we and our colleagues have observed several deficiencies when using RAPTOR. It may fail to include beads that lie over darker areas; e.g., when isolated cells are surrounded by resin or empty ice, it may select almost exclusively beads over the empty areas. In fact, the opposite is often desired for reconstructions from plastic-embedded material because the empty resin changes under the beam in ways that differ from regions with cellular material. RAPTOR provides no way to limit the area from which beads are selected. At high tilt beads may be added that are far from the area being reconstructed. It fails with large areas, such as from montaged tilt series, whose significant nonlinear distortions must be fit with local alignments (Mastronarde, 2007), because it can only fit all points to a single, global alignment. It often does not behave well for the cryoET data available to us from our colleagues in Boulder (see Section 2.3). Finally, its bead positions are significantly less accurate than those produced by Beadtrack, a point recently reported by Han et al. (2015) and on which we provide data in Section 2.3.

We thus decided to build on the existing approach of seeding then tracking. To do so, it was necessary to develop a procedure for generating a seed model automatically. The overall goal was to generate a model at least as good as the one a user would choose manually: it should consist of beads that will track well and are well distributed both in the plane and between two sides of a section (when they have been deposited on both sides). There should also not be many more beads than the number needed to achieve a reconstruction of adequate quality for the scientific question being pursued, because having excessive beads creates extra labor if the manual procedures of filling in missing points and checking deviant positions are going to be used. The procedure followed by the script Autofidseed is as follows:

1. The program Imodfindbeads is run to detect all beads on 3–7 views near 0° tilt (Fig. 1A). This program first cross-correlates the images with a model bead of the specified size. It analyzes a smoothed histogram of correlation peak strengths to find a dip between two modes, first searching for this motif in a highly smoothed histogram (e.g., blue curves in Fig. 1C and A), then over a more restricted range in less smoothed histograms. The location of the dip in the least smoothed histogram examined (arrows and red curves, Fig. 1C and D) is considered the best threshold between beads and non-beads. A raw correlation score is used instead of the normalized cross-correlation coefficient (CCC) because the CCC loses the intensity information and does not discriminate well between beads and weaker features with the same shape. Points above the threshold, or above a conservative fallback threshold if the histogram analysis fails, are aligned and summed to get a reference for iterating the correlation, peak search, and histogram analysis. The final selection of beads is thus based on correlations with an averaged bead, not the model. To alleviate the bias caused by beads over low-density backgrounds having much stronger correlations, the correlation peaks are divided into 4 groups based on background density and the histogram for each group is analyzed separately. A relationship between threshold and background density is then derived from the thresholds found in the successful analyses, provided that at least two succeed. The criterion for deciding whether any given point is a bead can thus be based on its background density. When there are few beads, the number of projections analyzed is increased to 7 to make this analysis more robust. Fig. 1A shows an example of beads selected by this process.

Please cite this article in press as: Mastronarde, D.N., Held, S.R. Automated tilt series alignment and tomographic reconstruction in IMOD. J. Struct. Biol. (2016), http://dx.doi.org/10.1016/j.jsb.2016.07.011