doi:10.1017/S1431927618004828
Microsc. Microanal. 24 (Suppl 1), 2018

A Complete Workflow for Cellular Tomography and Subtomogram Averaging in EMAN2

James M. Bell[1,2], Muyuan Chen[2], Adam Fluty[2], Steven J. Ludtke[2]

[1.] Graduate Program in Quantitative and Computational Biology, Baylor College of Medicine, Houston, Texas
[2.] Verna Marrs and McLean Department of Biochemistry and Molecular Biology, Baylor College of Medicine, Houston, Texas

EMAN2 was originally used primarily as a tool for single particle analysis [1,2]. Over recent years we have introduced tools for portions of the single particle tomography (SPT, or subtomogram averaging) pipeline [3]. It has become clear that to perform all of the necessary corrections and achieve optimal resolution, it is necessary to consider the complete process, starting with the tilt series movies. We now present a complete interactive pipeline for processing cryotomography data from cells or purified material, making use of Deep Learning technologies, Fourier reconstruction, and other techniques (Figure 1A).

Raw CryoET data consists of a tilt series of micrographs recorded as the specimen is rotated in the column. High frame-rate direct detection devices (DDD) permit each tilt to be separated into a sequence of very low dose movie frames. Our workflow spans the entire process from aligning movie frames within each tilt series image to the final process of CTF corrected subtomogram averaging. A single project may incorporate many tilt-series/tomograms, and all data at each step of the process is organized using standard naming conventions so data and metadata are preserved. While we provide a complete pipeline, data may be injected into the process at any stage, for example, reconstructed tomograms may be imported instead of tilt-series, but doing this may limit some processing options later in the pipeline.

Processing begins with alignment of the per-tilt movies. Once one or more tilt series are generated, users are able to perform iterative, fully-automated tilt series alignments with or without the presence of gold fiducials, including tilt-axis localization. Aligned tilt series are reconstructed using Fourier methods which, while they use considerable RAM, are extremely fast and can produce fewer artifacts than methods such as simple back-projection. This process is iterated, making use of high contrast landmarks in 3D and attempt to aligning these features within the tilt images (Figure 1B). The final result is an optimized 3-D reconstruction, as well as an aligned tilt series, which can be used for CTF corrected subtomogram averaging.

Targeted features within reconstructed tomograms can be localized using our semi-automated workflow based on convolutional neural networks [4]. Not only does this localize putative particles for subtomogram averaging, but it can also provide critical information about the number and distribution of various molecular species within a cellular domain. Once basic subtomogram averaging has been performed, 3-D particle locations can be used to extract data from the original tilt-series, which permits more precise CTF and geometrical corrections. We are also developing other CNN approaches for handling compositional and conformational variability. The resulting 3-D structures can then be processed using the wide range of tools already developed in EMAN2 for single particle analysis.

This new pipeline increases not only the speed of tomogram analysis but also improves the quality and resolution of the resulting 3-D structures, permitting more effective studies of the increasing numbers of cellular CryoET tilt series now emerging from current generation instruments.

References:

[1] G Tang, L Peng, P Baldwin, D Mann, W Jiang, I Rees, S Ludtke, Journal of Structural Biology **157** (2007) p. 38-46.
[2] J Bell, M Chen, P Baldwin, S Ludtke, Methods **100** (2016) p. 25-34.
[3] J Galaz-Montoya, C Hecksel, P Baldwin, E Wang, S Weaver, M Schmid, S Ludtke, W Chiu, Journal of Structural Biology **194** (2016) p. 383-394.
[4] M Chen, W Dai, S Sun, D Jonaschm C He, M Schmid, W Chiu, S Ludtke, Nature Methods **14** (2017) p. 983–985
[5] The authors acknowledge funding from the National Institute of General Medical Sciences (NIGMS, R01GM080139). This work is also supported by the Houston Area Molecular Biophysics Program (T32GM008280) NIH training fellowship through the Keck Center of the Gulf Coast Consortia.

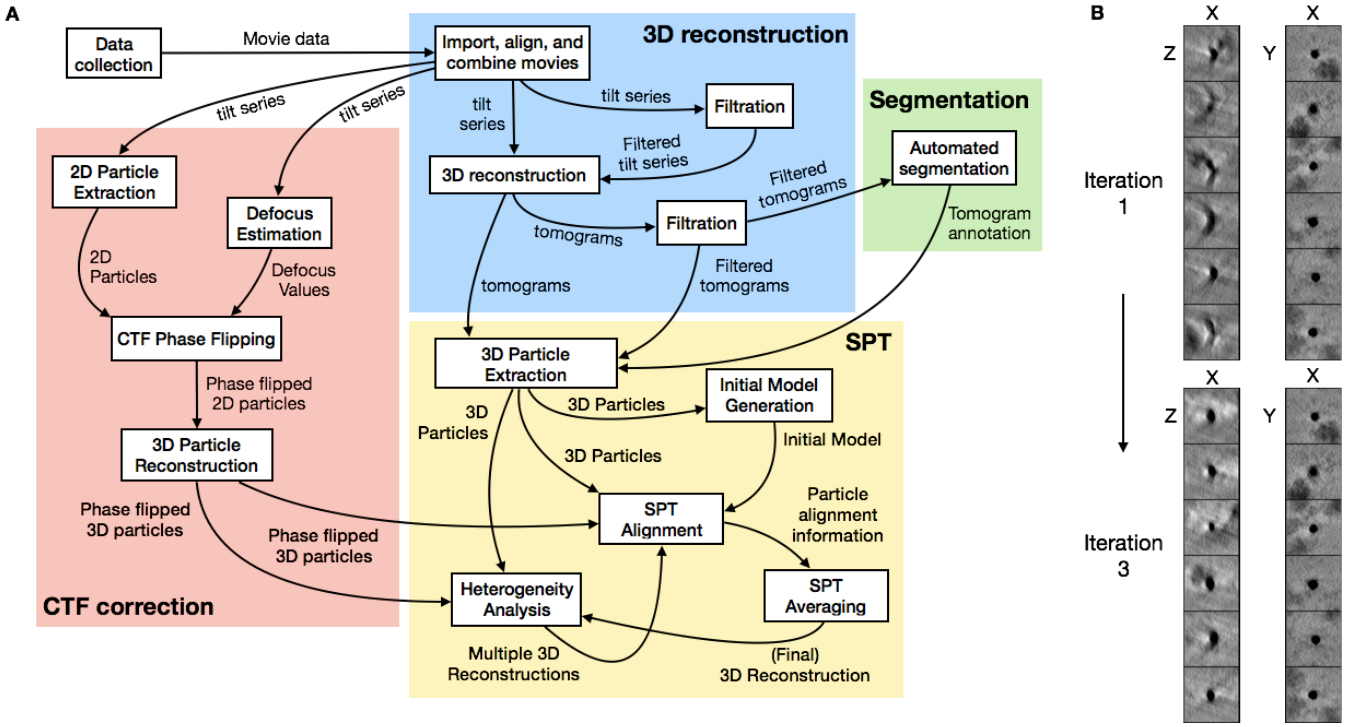

Figure 1A: visual overview of the complete tomography workflow available in EMAN2. We have expanded our workflow, offering tools for handling movie-mode images and performing iterative 3D reconstructions, automated segmentation, CTF correction, and subtomogram averaging.

Figure 1B: Landmark alignment optimization. Coarse tilt alignments yield features with characteristic smearing artifacts along their Z-axis. Our iterative alignment routine reduces these artifacts. Data was obtained from EMPIAR-10064.

https://doi.org/10.1017/S1431927618004828 Published online by Cambridge University Press