Interpreting Cytoskeletal Filaments in Cryo-Electron Tomograms with Shape-Constrained Deconvolution

Willy Wriggers[1], Manfred Auer[2], Salim Sazzed[3], Peter Scheible[3] and Jing He[3]

[1]Dept. of Mechanical & Aerospace Engineering and Institute of Biomedical Engineering, Old Dominion University, Norfolk, VA 23529, United States, [2]Molecular Biophysics and Integrated Bioimaging Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, United States, [3]Dept. of Computer Science, Old Dominion University, Norfolk, VA 23529, United States

The goal of EM imaging is to capture the 3D ultrastructure of supramolecular complexes in their native (organelles, cell, or tissue) environments. Cryo-Electron Tomography (cryo-ET) of unstained frozen-hydrated samples is widely used to achieve this aim. The data collection process in cryo-ET differs from that in single-particle cryo-EM. Instead of taking single images of multiple objects at random orientations, cryo-ET collects multiple images of the same object, with the specimen being tilted at different angles. Due to tilt angle limitations and limits on the acceptable total radiation dose, tomograms often display reconstruction artifacts, anisotropic and low resolution (40–50 Å), and a high level of noise. Specifically, orientational (missing wedge) artifacts can be prominent (diagonal streaks in the center panel of Fig. 1).

Our team studies hair cell stereocilia with a shaft region consisting of near-hexagonally packed actin filaments, and a less regular taper region (Fig. 1). Initially, we focused on the regular shaft region and developed an automated tracing method based on a 7-filament bundle template [1]. Due to the signal blending caused by missing wedge effects, we encountered difficulties using these methods in the taper region, where nearby filaments were not aligned with each other. Previous approaches by us and other groups have addressed the problem of tracing complex networks of actin filaments by "template convolution" (i.e., maximizing the local cross-correlation with a single shape template as a function of the template's direction at each voxel position [2]). Template convolution, however, requires significant separation between filaments to avoid the fusion of densities we observed. Therefore, we recently introduced in [3] a deconvolution approach that fully automates the tracing for "hard" cases that exhibit high noise and the problematic missing wedge-induced fusion of filamentous features. Instead of taking the (imperfect) experimental map at face value, the deconvolution seeks the "true" map (i.e., the most likely model— informed by the shape template—that explains the observed data; Fig. 1). The "true" map can subsequently be traced as described elsewhere [4]. The automatically traced filaments overlap well (within approximately 1 voxel) with a manual tracing [5]. Although we require a specific orientation of the template in a chosen volume, the approach is robust for angular variations of up to 25° from the main orientation [3].

Our original approach [3] re-framed the deconvolution into a constrained quadratic programming problem, which was executed with the proprietary SNOPT library [6] due to a lack of free alternatives. Consequently, the time of execution on large tomograms was inefficient (in the order of a week for Fig. 1). The difficulties were further compounded by the cost of the license and the complex software installation. In the current work we offer instead a novel implementation based on iterative Richardson Lucy (RL) deconvolution [7]. Assuming a Poisson noise model, RL deconvolution estimates the "true" density map by iteratively improving the log-likelihood. RL can be readily accelerated via GPU using 3D FFT and other CUDA libraries. This alone results in a noticeable acceleration in time for large tomograms. On our stock machine with a consumer grade GPU (CPU: Intel Core i7-4790 CPU @ 3.60GHz; GPU: Nvidia GeForce GTX 1050 Ti), we found that RL implementation (2000 iterations needed to give comparable results) was generally 42.4x faster than SNOPT. Additionally, utilizing a more advanced Nvidia Tesla P100 GPU, we achieved a gain of 147.8x over SNOPT. Therefore, we can now process an entire tomogram like that in Figure 1 in one hour instead of one week. We foresee additional acceleration through natural capabilities in micro-batching across multiple GPUs due to the sub-partitioning of the tomogram already inherent in our method.

Comparing the various tracing methods (manual, SNOPT, and RL) we can identify notable similarities using point cloud distance analysis [3], but without a quantitative ground truth for the tomograms, it is difficult to concretely determine and compare their accuracy. Thus, it is of great importance to utilize a simulated cryo-ET pipeline for

the generation of synthetic ground truths of complex cellular structures: We describe such representations of full cellular phantom structures in silico in more detail elsewhere [4].

Template-constrained deconvolution is a mathematically rigorous solution to the well-known problem of the missing wedge that has vexed cryo-ET for decades. Although we use templates in our modeling, the deconvolution seeks a model that best describes the experimentally observed data before performing any detection of filaments. Such a reconstruction of the experimental map is similar to the regularization proposed in cryo-ET more than two decades ago by Skoglund et al. [8], but instead of producing "the most featureless reconstruction" in the deconvolution (useful for denoising a map), we inform the reconstruction of the map with the known (or desired) template shape.

The use of a template is the key advantage over the earlier regularization: Our approach prescribes the geometric structure of the template onto the map T. We seek to extend the application to other shapes of cellular building blocks, such as membranes, microtubules, and storage vesicles. We have already tested filamentous sheet templates that could be used in a future membrane detection algorithm. Constrained deconvolution could also become an excellent complementary alternative to sub-tomogram averaging when averaging is not able to reduce missing-wedge artifacts due to the preferred orientation of the particles. Another benefit of RL is the possibility of stopping the deconvolution process for lower grade approximations early. Also, since the new RL implementation is of our own making, it can be freely disseminated as open source.
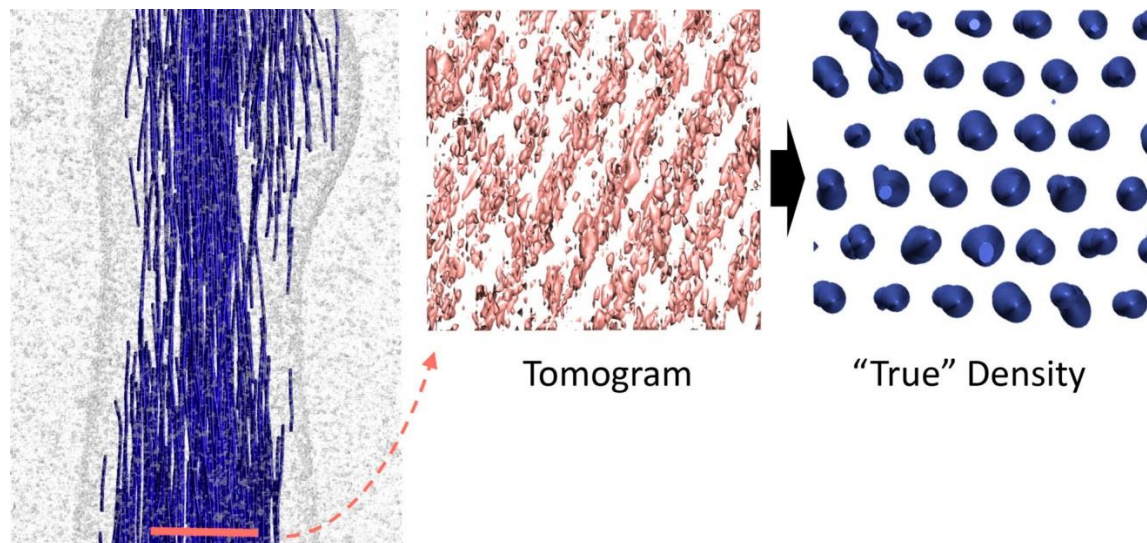

Figure 1. Figure 1. Tomogram (gray) of the hair cell stereocilium taper region [5] and a 30-voxel-thick cross section (red) showing missing wedge artifacts and noise at the boundary to the regular shaft region. We propose to reconstruct the map to its "true" density so that it is free of the artifacts and is suitable for automated tracing of actin filaments (blue).

References
[1] S Sazzed *et al*, Molecules, **23** (2018) p. 882.
[2] M Rusu *et al*, J. Struct. Biol., **178** (2012) p. 121.
[3] J Kovacs *et al*, J. Chem. Inf. Model., **60** (2020) p. 2626.
[4] S Sazzed *et al*, "Tracing Filaments in Simulated and Experimental 3D Cryo-Electron Tomography Maps Using a Fast Dynamic Programming Algorithm". Submitted to Microscopy & Microanalysis (August 1-5, 2021) Pittsburgh, PA.
[5] J Song *et al*, J. Struct. Biol., **210** (2020) p. 107461.
[6] PE Gill *et al*, SIAM Rev., **47** (2005) p. 99.
[7] M Ingaramo *et al*, Chem. Phys. Chem. **15** (2014) p. 794.
[8] U Skoglund *et al*, J. Struct. Biol., **117** (1996) p. 173.
[9] The work in this paper was supported in part by funding from the Batten Endowment at ODU and NIH R01-GM062968.