

Accelerating Discovery in 3D Microanalysis: Leveraging Open Source Software and Deskside High Performance Computing

Terry S. Yoo¹, Bradley C. Lowekamp¹, Oleg Kuybeda¹, Kedar Narayan², Gabriel A. Frank², Alberto Bartesaghi², Mario Borgnia², Sriram Subramaniam², Guillermo Sapiro³, and Michael J. Ackerman¹

¹ Office of High Performance Computing and Communications, National Library of Medicine, U.S. National Institutes of Health, Bethesda, Maryland, USA

² Laboratory of Cell Biology, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA

³ Electrical and Computer Engineering, Duke University, Durham, NC, USA

The recent decade has seen a dramatic elevation in the computing power affordably and routinely available to biological laboratories. Computing cores in servers, desk-side, and even laptop computers have doubled in number and capability on the order of every two years, making workstations today the rivals of supercomputers from the year 2000. 64-bit processors have substantially increased addressable main memory, and computational analysis can now keep pace with the growing size of datasets. Commensurately, the operation of the microscopes has become increasingly sophisticated, evolving from analog consoles to digital interfaces. These developments enable the automation of image collection, storage and quantitative analysis of the resulting data. Inexpensive storage and high bandwidth in digital networks promote the sharing of data and broad multidisciplinary interaction among research groups.

We present examples of how our collaboration between high performance computing developers at the National Library of Medicine and high-resolution electron microscopists at the National Cancer Institute are working to advance and develop new science. Our system designers combine innovative software and commercial computing hardware such as graphics processing units (GPUs) originally developed for entertainment to accelerate the solving of complex problems in microbiology. We concentrate much of our effort on the use of open source software, developed in part by our team or sponsored by our programs and repurposed for microanalysis [1].

Super-resolution through sub-volume averaging: In recent years, our group has been developing methods in transmission electron tomography to resolve glycoprotein complexes on the surfaces of viruses at the scale of a single nanometer. By accurately aligning and averaging very noisy sub-volumes we decode the structure of molecules at the edge of the resolution of our instruments. Starting with the early successes of our team, we have redesigned our software to use a combination of conventional code (written in MATLAB using the Parallel Toolbox) and highly parallel subroutines designed as NVIDIA CUDA kernels for singular value decomposition, programmed in C++, and run on multiple GPUs. The resulting architecture can generate results an order of magnitude faster than previously achieved using hardware that costs less than a tenth the price of earlier computing clusters. Results that previously took days can be achieved in hours using machines that comfortably fit in a laboratory budget [2].

Segmentation and visualization using open source software: Software originally designed for computer aided diagnosis and computer assisted surgery can be profitably applied to 3D microanalysis. We have adapted algorithms and filters from the open source Insight Toolkit (ITK) for the study of data from FIB-SEM data. Further, we have adapted 3DSlicer [3], a free open source medical image analysis console for quantitative research to 3D high-resolution microanalysis. We supplement these efforts by

reducing the programming burden through the introduction of SimpleITK, a scripting interface for ITK that provides easy access to hundreds of filters and algorithms inside 3DSlicer [4]. The result is a powerful and versatile platform for 3D microanalysis that leverages research by other disciplines.

We encourage a policy for software and data sharing that helps ensure reproducible results, empowers the entire field, and stimulates invention. It not just the quality of the data you can generate with your instrument, but the significance of the analysis that drives our science. The group that is the most adept and facile in the use of today's tools is the team that has the greatest advantage [5].

References:

- [1] TS Yoo, *et al.*, Visualizing cells and humans in 3D: biomedical image analysis at nanometer and meter scales. *IEEE Computer Graphics and Applications*, **32**(5)(Sep-Oct 2012), 39-49.
- [2] O Kuybeda, *et al.*, [A collaborative framework for 3D alignment and classification of heterogeneous subvolumes in cryo-electron tomography](#), *Journal of Structural Biology*, **181**(2) (2013), Academic Press, pp. 116-127.
- [3] A Federov, *et al.* 3DSlicer as an image computing platform for the quantitative imaging network, *Magnetic Resonance Imaging*, **30** (2012), DOI: 10.1016/j.mri.2012.05.001, pp. 1323-1341.
- [4] BC Lowekamp, DT Chen, L Ibanez, and D Blezek, The design of SimpleITK, *Frontiers in Neuroinformatics*, **7**(45) (2013), DOI: 10.3389/fninf.2013.00045, pp. 1-14.
- [5] SimpleITK was funded in part under the American Reinvestment and Recovery Act (ARRA) by NLM contract awards HHSN276201000488P and HHSN276201000490P. 3DSlicer is freely available from the National Alliance for Medical Image Computing (NA-MIC).

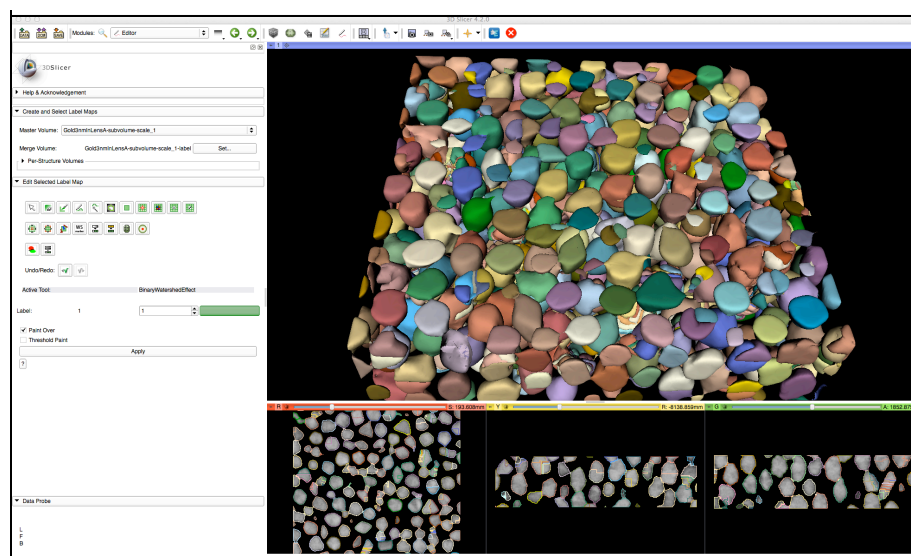


Figure 1. Screen shot of 3DSlicer (open source and available at www.slicer.org) and the IASEM extensions used to segment and label part of a dataset of 100nm gold particles embedded in resin. The volume was acquired with 5nm resolution collected with a FIB-SEM system. The segmentation uses 16 separate filters invoked through SimpleITK (www.simpleitk.org). The resulting full segmentation of the dataset delineates over 9,000 gold beads, represented in the figure above as a unique color overlay per bead over the original grayscale data, and volume rendered in the large window.