COMPUTATIONAL CHALLENGES IN HIGH-RESOLUTION CRYO-ELECTRON MICROSCOPY

Thesis by

Peter Anthony Leong

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2009

(Defended Aug 04, 2008)

© 2009

Peter Anthony Leong

All Rights Reserved

To God, who made all this possible

Acknowledgements

I would like to thank Prof. Grant Jensen for being an excellent advisor and role model for me during my time at Caltech. He has always shown tremendous understanding, kindness, and encouragement throughout, and supported me to the fullest extent in my research, career, and personal development — making this a wonderful life experience.

I am also very indebted to Drs. Bernard Heymann and Andrew Rawlinson. Bernard, who was a mentor to me when I first joined the lab, taught me much about computer hardware and software. Andy, who worked closely with me during the middle of my thesis work, helped me greatly in our discussions about the mathematics and physics related to our research work. Their mentoring has been very significant in my development as a scientist. In addition, I would also like to thank all my other lab mates, both past and present, who have been wonderful colleagues and friends. This thesis work could not have been completed without their help.

I would also like to thank my thesis and candidacy committee members Profs. Scott Fraser, Douglas Rees, Brent Fultz, Robert Phillips, and Z. Hong Zhou. Their advice and feedback about my research projects and about academics in general have been extremely helpful.

I would also like to thank my collaborators Profs. Hong Zhou, Wen Jiang, and Nikolaus Grigorieff, and their respective lab members, especially Drs. Xuekui Yu and Weimin Wu, for supporting me and helping in the completion of my thesis work.

Lastly and most importantly, I would like to thank my family, especially my parents, for all the love, support, and encouragement they have always shown me.

Abstract

To avoid the challenges of crystallization and the size limitations of NMR, it has long been hoped that single-particle cryo-electron microscopy (cryo-EM) would eventually yield atomically interpretable reconstructions. For the most favorable class of specimens (large icosahedral viruses), two of the key obstacles are the large computational requirements of high-resolution reconstructions and the curvature of the Ewald sphere, which leads to a breakdown of the projection theorem used by conventional 3D reconstruction programs. Here, two solutions to these obstacles are presented.

First, a simple distributed processing system named Peach was developed to meet the rising computational demands of modern structural biology (and other) laboratories without additional expense by using existing hardware resources more efficiently. A central server distributes jobs to idle workstations in such a way that each computer is used maximally, but without disturbing intermittent interactive users. As compared to other distributed systems, Peach is simple, easy to install, easy to administer, easy to use, scalable, and robust. While it was designed to queue and distribute large numbers of small tasks to participating computers, it can also be used to send single jobs automatically to the fastest currently available computer and/or survey the activity of an entire laboratory's computers. Tests of robustness and scalability are reported, as are three specific cryo-EM applications where Peach enabled projects that would not otherwise have been feasible without an expensive, dedicated cluster.

Second, an iterative refinement reconstruction algorithm, *Prec*, is described that overcomes the curvature of the Ewald sphere resolution limitation by averaging information from images recorded from different points of view, as are present in typical micrographs. *Prec* was implemented in the popular software packages IMIRS, EMAN, and Bsoft. In preliminary tests with both simple and multi-slice simulated images, *Prec* overcame the curvature problem even in the presence of noise. *Prec* was then used to refine the three recently published, ~ 4 Å resolution, icosahedral virus reconstructions from experimental cryo-EM images, but unfortunately no significant improvements in resolution were realized. Further simulations showed that limitations other than the Ewald sphere curvature problem must still be dominant in these experimental studies.

Table of Contents

List of Figures and Tables