Professor Peter Doerschuk project:

5). Software engineering for a distributed memory high performance image processing system

Electron microscopy is a critically important method for 3-D visualization of nanoscale biological machines such as ribosomes and viruses. Electron microscopy of snap-frozen specimens, so-called cryoEM, provides fundamentally tomographic projection information. Therefore, the process of computing a 3-D image of the biological object is similar to the process used to compute a 3-D image of a patient in a medical x-ray computed tomography scanner. However, the cryoEM problem is more challenging because the signal to noise ratio is lower, the orientation of the tomographic projections is not known, and the microscope measures the tomographic projections modified by a so-called Contrast Transfer Function, which is a linear system with zeros in the spatial frequency band of interest. With graduate students and collaborators at The Scripps Research Institute, I (Peter Doerschuk) have been developing algorithms and parallel software implementations of the algorithms for a variety of specific cryoEM problems since about 1999. As of Fall 2007, there are four Ph.D. students working on different aspects of cryoEM who have interests in signal and image processing, software engineering, or computational biology. I am proposing several M.Eng. projects that concern different aspects of computational cryoEM. The content of the projects can be rearranged in order to accommodate the interests of individual students or of groups of students. Specific M.Eng. Contribution: We have algorithms and parallel software that successfully computes 3-D visualizations of viruses. There are actually four somewhat different codes which each solve different specific cryoEM problems and which are all eventually related to a code developed in the software engineering masters degree thesis of Yili Zheng who is now working on his Ph.D. thesis. The goal of this M.Eng. project is to work on the parallel C code. Some of the improvements are software-centric, e.g., interfacing additional image format capabilities and replacing code for standard numerical operations (e.g., solving systems of linear equations) that is not public domain. Other of the improvements are on the boundary between algorithm and software, e.g., altering the indexing in order to achieve stride-one behavior in the cache, replacing redundant calculations by more sophisticated control structures, and integration of BLAS calls to perform the fundamental linear algebra operations. Depending on the student's interests, there is also the possibility of implementing algorithmic changes, e.g., implementing additional sets of basis functions. Peter Doerschuk, who wrote the original serial program, and Yili Zheng, who developed multiple methods for parallelization, continues to be interested and involved with these codes and so even if only one student rather than a group of students was working on the project, the student would not be working alone. Based on the segmentations for each group of rats.

Contact information: Peter Doerschuk (pd83@cornell.edu)