CODA:
Looking Toward the Future

Some open problems in tomography
Petascale computing meets reality
Ongoing Projects/Wish Lists

- Better means of data and code sharing
- Extending TxBR to make better use of cluster and grid computing
- Improved AI for tracking, feature recognition and segmentation
- Reconstruction methods more attuned to the nature of the instrument
- Improvement of reconstruction algorithms
Better Means of Data and Code Sharing

- Problems with older image data in repositories like CCDB
- Outmoded image formats and no pointers to programs to translate or manipulate data
- Data obtained by latest instrumentation or collection techniques is slow to get posted
- Candes has proposed a standard for research codes posted on the net
- Inclusion of test data with all I-published code
- Standardization of image formats
- Divergence of purpose: mathematicians and computer scientists need to work with latest acquisition methods; biologists want biomedical relevance
Extending TxBR to Make Better Use of Cluster and Grid Computing

- TxBR was designed to bring advanced methods to bear on very large data sets
- Parallel implementation is a bonus because conscious effort was made to make the math as general as possible
- Packaging TxBR to take advantage of GPU boards
- Do we bring the data to the code or the code to the data
- Code is orders of magnitude smaller than data
- Processing code could be part of file headers
- Look at biological examples: Where is the code? Where is the data?
Improved AI for Tracking, Feature Recognition and Segmentation

• Tracking and segmentation are the labor-intensive parts of the production cycle at NCMIR
• 1 day to track markers for alignment, 0.5 days to align, filter and backproject, 5 days to mark surfaces around structures of interest
• AI methods are designed for serial computers and often perform relatively poorly
• But the brain is massively parallel and performs the same tasks very well
• Design of clever algorithms to move the burden of the combinatorics to parallel computation
Reconstruction Methods More Attuned to the Nature of the Instrument

- First EM tomography codes modeled the instrument as an X-ray machine
- Compensations for differential rotation and magnification added (IMOD)
- Compensations for transfer function (Single-Particle Codes)
- New modes of EM: dark field, scattering
- Forward problem can be modeled as a Fourier integral operator, or with scattering, an ordered exponential transform (think of a path integral in operator space)
- Inverse problem is of considerable mathematical and computational difficulty
- Petascale computing necessary for even small data sets
Improvement of Reconstruction Algorithms

• Better filtration methods
• Removing noise and artifact
  – Three kinds of artifact, discretization, limited angle, and intrinsic sampling bias
  – Cross validation
• Regularization methods
• Dewarping and montaging
  – Sample warping
  – Mismatch across boundaries due to electron trajectory mismatch
• Compensation for defocus