Available online at www.sciencedirect.com

ScienceDirect

Future Generation Computer Systems 23 (2007) 435–446

FGCS
FUTURE GENERATION COMPUTER SYSTEMS

www.elsevier.com/locate/fgcs

Electron tomography of complex biological specimens on the Grid

José-Jesús Fernández[a], Inmaculada García[a], Jose-María Carazo[b], Roberto Marabini[c,*]

[a] Department of Computer Architecture and Electronics, University of Almería, 04120 Almería, Spain
[b] Centro Nacional Biotecnologia - CSIC, Campus Universidad Autonoma, Cantoblanco, 28049 Madrid, Spain
[c] Escuela Politecnica Superior, Universidad Autonoma de Madrid, Cantoblanco, 28049 Madrid, Spain

Received 14 December 2005; received in revised form 21 April 2006; accepted 15 July 2006
Available online 22 August 2006

Abstract

Electron tomography allows elucidation of the three-dimensional structure of large complex biological specimens at molecular resolution. To achieve such resolution levels, sophisticated algorithms for tomographic reconstruction are needed. Iterative algebraic algorithms yield high quality reconstructions, but they are computationally expensive and high performance techniques are needed to exploit them in practice. We present here a grid computing approach for tomographic reconstruction of large biological specimens. The approach is based on the computational Single-Program-Multiple-Data model, which basically decomposes the global problem into a number of independent 3D reconstruction subproblems. New performance metrics and job submission policies are proposed here that could be of general interest in the field of Grid Computing. We have evaluated this approach on the grid hosted by the European EGEE (Enabling Grids for E-sciencE) project. The influence of the problem size and the parallelism grain has been thoroughly analyzed. Our results demonstrate that the grid is better suited for large reconstructions, as currently needed in electron tomography. To fully exploit the potential of computational grids, the global problem should be decomposed into an adequate number of subdomains.

Keywords: Grid computing; Electron tomography; Electron microscopy; Structural biology; Reconstruction algorithms

1. Introduction

Knowledge of three-dimensional (3D) structure of biological specimens is critical to understanding their function at all levels of detail [1]. Electron tomography (ET) combines the principles of tomographic imaging with the best possible preservation of the samples to elucidate the structure of large complex biological specimens at molecular resolution [2–4]. This resolution allows the identification of single macromolecules and molecular interactions in their native cellular context [5], which is critical to understand the cellular function [1]. ET has already made possible the visualization of the structure of large complex viruses, organelles and even whole cells [6–12]. In ET, a set of images from a single individual specimen is acquired at different orientations. Those images represent projections, or in other words “radiographs”, of the specimen. From those images, a 3D reconstruction can be derived by means of tomographic reconstruction algorithms. ET is very similar to computer-assisted tomography in medicine.

Rigorous structural analyses require that image acquisition and reconstruction introduce as little noise and artifacts as possible at the spatial scales of interest, for a proper interpretation and measurement of the structural features. As a consequence of the resolution needs, ET of complex biological specimens requires large projection images (typically 1024 × 1024 pixels or larger). ET on this scale yields large reconstruction files and requires an extensive use of computational resources and considerable processing time. High Performance Computing has turned out to be key to address such computational requirements [13,14], thereby allowing determination of the 3D structure of large volumes in reasonable computation time.

Weighted back-projection (WBP) [15] has been one of the most popular methods in ET due to its relatively low computational demands. Series expansion reconstruction methods [16] constitute one of the main alternatives to WBP in image reconstruction. Their iterative nature provides an implicit

---

* Corresponding author. Tel.: +34 91 585 4510; fax: +34 91 585 4506.
E-mail addresses: jose@ace.ual.es (J.-J. Fernández), roberto@cnb.uam.es (R. Marabini).

0167-739X/$ - see front matter
doi:10.1016/j.future.2006.07.010

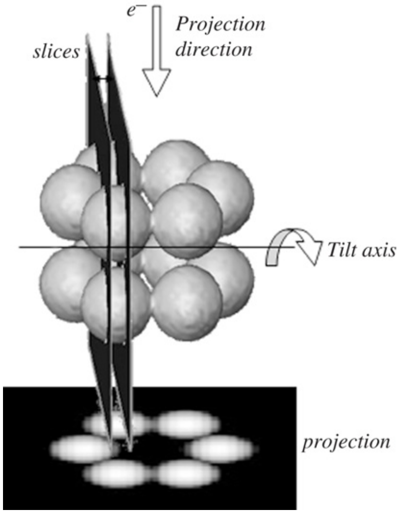

Fig. 1. Single-tilt axis data acquisition geometry. The specimen is imaged in the microscope by tilting it over a range typically +/−60 or 70 degrees in small tilt increments. As a result, a set of projection images needed to determine the structure is collected. In the figure, a slice perpendicular to the tilt axis is sketched.

regularization mechanism which makes them better suited for the noisy conditions found in ET [13,17]. Despite their potential advantages compared to WBP [13,17,18], these methods have not been extensively used in ET due to their high computational costs. However, they are lately receiving increased interest [19–21]. Parallel and distributed computing [22] emerge as a crucial tool to make the use of iterative methods affordable. Efficient parallel solutions for iterative reconstruction algorithms have been proposed [13,14,23–25] and their suitability for grid computing has also been analyzed and stated [26,27].

In this work, we present a grid computing solution for iterative reconstruction methods applied to ET. The approach is based on one-dimensional domain decomposition and aims to exploit the high-performance, high throughput capabilities of computational grids. Our approach has been tested and evaluated on the grid hosted under the European EGEE (Enabling Grids for E-sciencE) project [28,29].

2. Reconstruction methods in electron tomography

In electron tomography, the so-called single-axis tilting is the data collection geometry of choice to record the set of projection images required to compute the 3D reconstruction. For the collection of a standard single-axis tilt series, a single specimen is tilted over a range +/−60 or 70 degrees in small tilt increments (1–2 degrees), and an image of the same object area is recorded at each tilt angle via, usually, CCD cameras (Fig. 1). Electron tomographic datasets usually range from 60 to 150 images. Due to the resolution requirements, the image size typically ranges between 512 × 512 and 1024 × 1024 pixels.

The computation of a distortion-free 3D reconstruction from a single-axis tilt series would require that a dataset from the full tilt range (+/−90 degrees) is available. Due to physical limitations of electron microscopes, the tilt range is limited and, as a result, tilt series have a wedge of missing data corresponding to the uncovered angular range. This limitation, which is not present in medical tomography, causes distortions in 3D reconstructions: elongation along the electron beam direction and anisotropic resolution, so that features oriented perpendicular to the tilt axis tend to blur. On the other hand, the use of minimum electron doses in imaging the specimen is a requirement to achieve high resolution. Electron dose must be kept within tolerable limits to prevent radiation damage that may erase fine details of the structure. As a result, the signal-to-noise ratio in the projection images taken with the microscope is extremely low, typically lower than 0.1. Therefore, ET requires a method of "3D reconstruction from projections" able to cope with limited angle conditions and extremely low signal-to-noise ratios of the projection images.

Currently, the standard method in the field is the well-known Weighted Backprojection (WBP). Roughly, the method distributes the known specimen mass present in projection images evenly over computed backprojection rays. In this way, specimen mass is projected back into a reconstruction volume (a detailed description is available in [15]). When this process is repeated for a series of projection images recorded from different tilt angles, backprojection rays from the different images intersect and reinforce each other at the points where mass is found in the original structure. Therefore, the 3D mass of the specimen is reconstructed (Fig. 2, left). The relevance of WBP in ET mainly stems from its linearity and its computational simplicity. However, it has serious disadvantages for this particular problem, as (i) the results may be strongly affected by limited tilt angle data, and (ii) WBP does not implicitly take into account the noise conditions.

3. Iterative reconstruction methods

Series expansion reconstruction methods assume that the 3D object or function f to be reconstructed can be approximated by a linear combination of a finite set of known and fixed basis functions b_j

$$f(r, \phi_1, \phi_2) \approx \sum_{j=1}^{J} x_j b_j(r, \phi_1, \phi_2) \tag{1}$$

(where (r, ϕ_1, ϕ_2) are spherical coordinates), and that the aim is to estimate the unknowns x_j. These methods also assume an image formation model where the measurements depend linearly on the object in such a way that:

$$y_i \approx \sum_{j=1}^{J} l_{i,j} x_j \tag{2}$$

where y_i denotes the ith measurement of f and $l_{i,j}$ the value of the ith projection of the jth basis function.

Under those assumptions, the image reconstruction problem can be modeled as the inverse problem of estimating the x_j's from the y_i's by solving the system of linear equations given by Eq. (2). Such systems of equations are typically solved by means of iterative methods. One of the most popular iterative