7

Algorithms for Three-dimensional Reconstruction From the Imperfect Projection Data Provided by Electron Microscopy

Jose-Maria Carazo, Gabor T. Herman, Carlos O. S. Sorzano and Roberto Marabini

Jose-Maria Carazo • Centro Nacional de Biotecnología (CSIC), Universidad Autónoma, 28049 Cantoblanco, Madrid, Spain

Gabor T. Herman • Department of Computer Science, The Graduate Center, City University of New York, New York, NY 10016, USA

Carlos O. S. Sorzano • Escuela Politécnica Superior, Universidad San Pablo-CEU, 28668 Boadilla del Monte, Madrid, Spain

Roberto Marabini • Escuela Politécnica Superior, Universidad Autónoma, 28049 Cantoblanco, Madrid, Spain

1. INTRODUCTION

Since the 1970s, it has become increasingly evident that transmission electron microscopy (TEM) images of typical thin biological specimens carry a large amount of information on 3D macromolecular structure. It has been shown many times how the information contained in a set of TEM images (2D signals) can determine a useful estimate of the 3D structure under study.

In its most general form, the 3D reconstruction problem in TEM can be defined by the statement: given a collection of projection images (2D data) g, determine the 3D structure f that produced the images g. This problem has to be solved under conditions in which the image data, as well as the information about the geometry of data collection that relates g to f, are imperfect; in particular, both the gray level information in the images and the information regarding the direction of projections are corrupted by substantial noise. We are interested in knowing under which conditions g is adequate for producing an $\hat{f}$ that is close to f in some sense. For experimental reasons, we always have three basic limitations on the collection of the image data set g: the image gray level noise, the imperfect information about the data collection geometry and, finally, the restriction to a finite number of images. Can we do something to ameliorate the situation? Can we qualify our confidence in the reproducibility of the reconstructed structures (in other words, do small changes in g produce radical changes in the estimate $\hat{f}$?). Formulated in this way, it becomes clear that the topics covered under 3D reconstruction in TEM belong to the broad class of signal-recovery/inverse problems. It is within this framework that we discuss how the physical limits directly affect the fidelity of the estimated 3D structures produced by various reconstruction algorithms.

2. OVERVIEW OF OUR APPROACH

In this section, we present the concepts that we consider fundamental to a careful discussion of our topic. We illustrate these concepts as we go along on the relatively simple problem of recovering a 2D structure from its 1D images.

In the area of inverse problems, it is assumed that we have some understanding of the data collection process. Usually we describe it by some mathematical idealization, combined with a description of how the actual data collection process differs from the ideal one. For example, in the field of (2D) image reconstruction from (1D) projections, the ideal data consist of the line integrals, for all lines, of the image f to be reconstructed. In practice, we can have data corresponding to only finitely many lines, and the data are likely to be contaminated with noise (see Fig. 1). To approximate f from such data, there are two basic approaches that we call transform methods and series expansion methods, respectively.