3-D Reconstruction of 2-D Crystals in Real Space

Roberto Marabini, Carlos O. S. Sorzano, Samuel Matej, José J. Fernández, José M. Carazo, and Gabor T. Herman

***Abstract*—A new algorithm for three-dimensional reconstruction of two-dimensional crystals from projections is presented, and its applicability to biological macromolecules imaged using transmission electron microscopy (TEM) is investigated. Its main departures from the traditional approach is that it works in real space, rather than in Fourier space, and it is iterative. This has the advantage of making it convenient to introduce additional constraints (such as the support of the function to be reconstructed, which may be known from alternative measurements) and has the potential of more accurately modeling the TEM image formation process. Phantom experiments indicate the superiority of the new approach even without the introduction of constraints in addition to the projection data.**

***Index Terms*—3-D reconstruction, crystals, electron microscopy, image reconstruction, projections.**

I. Introduction

DETERMINATION of the structure of large macromolecular assemblies is a crucial subject in current biochemical research. Many basic biological processes, including DNA metabolism, photosynthesis, protein synthesis, and viral assembly, require the concerted action of a large number of macromolecules. Understanding their three-dimensional (3-D) organization in as much detail as possible is central to the interpretation of their function. Interest in the structure of biological molecules dates to the first attempts at X-ray diffraction from protein crystals. However, it was the invention of the electron microscope that made possible direct imaging of biological structures at a macromolecular level and, more recently, even at atomic resolution [1]–[6].

The analysis by transmission electron microscopy (TEM) of biological material is inconvenienced by its sensitivity to electron radiation. In order to minimize the damage caused by the radiation, the electron dose is kept low and consequently the signal-to-noise ratio of the images is poor.

Image processing methods have been developed to counter this problem. Although, in theory, these methods can be applied to a TEM image of any object, they are usually most powerful for objects in which subunits are arranged in a regular manner, such as two-dimensional (2-D) crystals. When processing images of such objects, Fourier transformation is usually employed [7], [8]. This is due to the fact that their Fourier transform (FT) is zero valued except at isolated points. This property often enables a convenient manipulation of the data. 2-D crystals are currently the only way to reach atomic resolution using electron microscopy.

In this work an iterative method for performing 3-D reconstruction of 2-D crystals in real space is proposed. We see two potential advantages of this approach over the traditional methods in Fourier space. First, it avoids interpolations in the Fourier domain that often generate artifacts (especially when the number of projections is small). Second, it allows the convenient incorporation of information available from other microscopies [9]–[11] and from adequate modeling of either the sample [12] or the imaging device [13].

The paper is divided as follows. Section II summarizes the traditional method for 3-D reconstruction of 2-D crystals and introduces the background needed for understanding our new reconstruction algorithm. Section III describes the new algorithm. Section IV discusses some details of implementation. Section V presents our strategy for comparing reconstruction methods and reports the results of our comparisons. Finally, Section VI discusses the results.

II. Theoretical Background

A. Basic Definitions

We start this section by stating our notation and defining rigorously the terms *2-D crystal, crystal vector,* and *unit cell.* We adopt the common mathematical terminology that functions assign *values* to points (these values can be real numbers or complex numbers) and that the *support* of a function is the (closure of the) set of points at which its value is not zero. We denote by bold letters, such as $\mathbf{c}$, J-dimensional row *vectors* (where J is some positive integer, in this paper usually 2 or 3), as in $\mathbf{c} = (c_1, \ldots, c_J)$. While the mathematical derivations below are not influenced by such interpretations, it is worthwhile to bear in mind that a vector can be used to represent both a point relative to a coordinate system (in space in case $J = 3$, and in the plane in case $J = 2$) and also the directed line segment from the origin (the zero vector) to that point. On this basis, we will refer to a vector as a *point* whenever such usage matches better the geometrical interpretation.

A 3-D function f is a *2-D crystal* if there are two linearly independent vectors $\mathbf{a} = (a_x, a_y, 0)$ and $\mathbf{b} = (b_x, b_y, 0)$, called the *crystal vectors,* such that

$$f(\mathbf{r} + m\mathbf{a} + n\mathbf{b}) = f(\mathbf{r}) \tag{1}$$

Manuscript received July 9, 2003; revised October 3, 2002. This work was supported in part by the Comisión Interministerial de Ciencia y Tecnología (Spain) and the National Institutes of Health (USA) through Grants P.N. Biotec. BIO 98–0761, BIO 2001–1237, TIC99–0361, TIC2002–00228, and HL70472.

R. Marabini is with the Escuela Politécnica Superior, Universidad Autónoma de Madrid, 28049 Madrid, Spain,

C. O. S. Sorzano is with the Biomedical Imaging Group (EPFL), CH-1015 Lausanne, Switzerland.

S. Matej is with the Department of Radiology, University of Pennsylvania, Philadelphia, PA 19104 USA.

J. J. Fernández is with the Departamento de Arquitectura de Computadores y Electrónica, Universidad de Almería, 04120 Almería, Spain.

J. M. Carazo is with the Centro Nacional de Biotecnología-CSIC, Campus Universidad Autónoma de Madrid, 28049 Madrid, Spain.

G. T. Herman is with the Department of Computer Science, Graduate Center, City University of New York, New York, NY 10016 USA.

Digital Object Identifier 10.1109/TIP.2003.822620

for all points $\mathbf{r}$ and for all integers m and n, and there is a positive number $z_{\max}$ such that $f(x,y,z) = 0$ if $|z| \geq z_{\max}$. Crystals are ordered structures that can be described as a repetition of identical *unit cells*. For the 3-D function f defined above, a possible unit cell is the set of points of the form $(ua_x + vb_x, ua_y + vb_y, z)$ such that $-1/2 < u \leq 1/2$, $-1/2 < v \leq 1/2$ and $-z_{\max} < z < z_{\max}$. We call this particular unit cell the *reference unit cell*. We use the word *motif* to denote the function that coincides with the 2-D crystal within the reference unit cell and is zero-valued outside the reference unit cell. Thus, a 2-D crystal can be thought of as consisting of regular repetitions of its motif.

B. Outline of the Fourier-Space Method of Crystal Reconstruction

The 3-D reconstruction process starts by obtaining a number of projections of the object at various angles and calculating their Fourier transforms (FTs). In the case of 2-D crystals, the projections can be described as the convolution of a *bidimensional Shah*[1] function (usually called the *real lattice*) and a 2-D motif. The FT of a Shah function is another Shah function (usually called the *reciprocal lattice*). Therefore, the projection FT will be the product of a Shah function and the FT of the 2-D motif. Even with imperfect real specimens the structural information in the FT is concentrated at the points of the reciprocal lattice, usually called *reflections* or *spots*. Once we have the FT's of all the projection images, the next step is to calculate the specimen 3-D FT by applying the central section theorem.[2] As we are working with crystals that repeat themselves only in two dimensions, the crystal 3-D FT is not a collection of spots, but is a collection of lines parallel to the Z-axis (that is, perpendicular to the crystal plane). These lines are called lattice lines and their intersections with the central sections are the spots whose values are given by the projection FTs. Since the specimen support is finite along the z direction, the values of the FT along the lattice lines change relatively smoothly and so, if the number of different projections is high enough, the specimen 3-D FT can be recovered.

C. 2-D Processing

Before starting the 3-D reconstruction, the 2-D projections obtained in the microscope need to be enhanced to increase their signal-to-noise ratio. We use the suite of programs developed at the *Medical Research Council* (Cambridge) [15] to obtain this goal. A detailed discussion is beyond the scope of this article; a description is provided by [16]. In summary, the programs perform two main steps. The first one is a filtration in Fourier space: since the FT of a perfect crystal projection should be confined to isolated points, every pixel not in a vicinity of these points can be forced to have the value zero. The second step consists of *unbending* or refining the crystal. Biological crystals are seldom perfect: they are limited in space, and present disorders. A repositioning of the individual unit cells over a perfect lattice (obtained by re-interpolation from the original image) usually improves the high resolution terms of the FT. In the MRC suite, the main program involved in the first step is called mmbox [7] and that in the second step is called unbend [17]. In our experiments that are reported below, lattice refinement was carried out using the X-windows based graphical environment SPECTRA [18]. Lattice distortion correction was performed by the methods described in [7]. In most cases two rounds of lattice unbending gave the best results (i.e., with the maximum number of high resolution spots in the projection FT).

D. 3-D Processing

The FT of a projection is stored in the MRC programs by a collection of isolated spots, each one of which has a complex value assigned to it that is represented by an "amplitude" and a "phase." To make the FT's of different projections consistent with each other, the amplitudes of the spots need to be scaled and the phases of the spots need to be shifted. The differences in amplitude between projections are due to the different number of unit cells contained in each image and the differences during the staining, development and/or digitization of the micrographs. The differences in phase are related to shifts in real space with respect to a common coordinate system. The scale factors and phase shifts that must be applied to each of the central sections need to be identified prior to creating the 3-D FT. In the MRC suite, the program ORIGTILT [13] performs these actions. The traditional approach and our new algorithm differ in how they process the *output* of ORIGTILT.

The traditional approach combines the projections in Fourier space to obtain a nonuniformly sampled 3-D FT. From this a uniformly sampled 3-D FT is obtained by 1-D interpolation along the lattice lines. The interpolation is made by a least-squares algorithm [19]. It is recommended that the experimental data be obtained at a third or fourth of the desired sampling rate in Fourier space, which is the inverse of the specimen thickness (according to our definition of a 2-D crystal, this thickness is never greater than $2z_{\max}$).

E. Algebraic Reconstruction Techniques

Several variants of the 3-D reconstruction algorithms known as Algebraic Reconstruction Techniques (ART) have been extensively tested in our laboratory for a variety of situations [20]–[24]. The conclusion we reach is that ART produces high-quality reconstructions and is superior, for a large number of tasks, to other well-known algorithms for 3-D reconstruction. Therefore, ART (and in particular a variant known as ART with blobs [23]) was selected as the core method for our developments for crystals. In this subsection we summarize our successful implementation of ART with blobs for single particles. The proposed modification needed to handle crystalline specimens is reviewed in Section III.

Algebraic reconstruction techniques are series expansion methods [25], i.e., they assume that a volume f can be approximated by a linear combination of a finite set of known basis

[1]A simple example of a bidimensional Shah function is: $^2\mathrm{III}(x,y) = \sum_{m=-\infty}^{\infty}\sum_{n=-\infty}^{\infty}\delta(x-m)\delta(y-n)$. See [14] for further details.

[2]The central section theorem states that the FT of any projection p of a volume f is equal to the restriction to one plane through the origin of the FT of f. This plane is often referred to as a *central section*.

functions b_j, each one of which is just the same function b shifted to one of J grid points (denoted by $\mathbf{g}_j$), as in

$$f(\mathbf{r}) \approx \sum_{j=1}^{J} c_j b_j(\mathbf{r}) = \sum_{j=1}^{J} c_j b(\mathbf{r} - \mathbf{g}_j) \tag{2}$$

and the task of the algorithm is to estimate the unknown coefficients c_j. A consequence of (2) is an image formation model of the form

$$y_i \approx \sum_{j=1}^{J} l_{i,j} c_j \tag{3}$$

where y_i is the ith measurement of the volume to be reconstructed (that is, a pixel value in the experimental data) and $l_{i,j}$ is the corresponding line integral of the basis function b_j. The values y_i and c_j form an PQ-dimensional vector and a J-dimensional vector, respectively (which we will denote by $\mathbf{y}$ and $\mathbf{c}$), where Q is the number of projections and P is the number of pixels per projection. The y_i elements are arranged in such a way that all pixels belonging to a projection are consecutive.

Conceptually, ART can be described as an algorithm that starts with a initial *trial volume* and iteratively modifies it depending on the differences between the measured projection data and the corresponding computer-generated projections of the current trial volume (see Fig. 1). For a review on ART, see [26].

The particular variant of ART that we use as the basis of our algorithm operates as follows. Starting with a J-dimensional zero vector for the estimate of $\mathbf{c}$, we update this estimate of $\mathbf{c}$ iteratively. In an iterative step, we make use of data from one projection only; we repeatedly cycle through all the projections in the complete algorithm. The update of the estimate of $\mathbf{c}$ is done in a computationally efficient manner so as to reduce the discrepancy between the measured data provided by the TEM projection used in that iterative step and the matching *pseudo projection data* that is obtained from the current estimate of $\mathbf{c}$ using the right-hand side of (3).

To be exact, the kth iterative step is

$$\mathbf{c}^{(k+1)} = \mathbf{c}^{(k)} + \lambda \sum_{i=(q-1)P+1}^{qP} \frac{(y_i - \langle \mathbf{l}_i, \mathbf{c}^{(k)} \rangle)}{\langle \mathbf{l}_i, \mathbf{l}_i \rangle} \mathbf{l}_i,$$
$$\text{for } q = [k(\mathrm{mod}\, Q)] + 1 \tag{4}$$

where λ is a real number called the *relaxation parameter* that controls the magnitude of each update, $\mathbf{l}_i$ is the J-dimensional vector whose jth component is $l_{i,j}$, and $\langle \cdot, \cdot \rangle$ denotes the inner product (dot product) between two vectors. The expression $[k(\mathrm{mod}\, Q)] + 1$ means that in the kth iterative step we use the projection indexed by the remainder of the division (k/Q) plus 1. A more careful description of this implementation, including convergence properties, is given in [27].

Following [28], we do not use voxels as the basis functions, but spherically symmetric functions that are not only spatially limited but can also be chosen to be smooth. They are the generalized Kaiser–Bessel window functions, also known as *blobs*. The property of blobs that makes them very promising for the formulation of many algorithms is that they are smooth both in the real and in the Fourier domain. The choice of the grid (defined as the set of locations of the centers of the blobs) influences the quality of the reconstruction. It was found that grids different from the cubic one are preferable, in particular the use of the so-called body-centered cubic grid (BCC, obtained by interlacing two cubic grids in such a way that the points of either grid are central to eight points forming a cube in the other grid) is recommended [29]. However, this recommendation is based on theoretical and experimental results for volumes that are not crystalline, it has to be modified for crystalline volumes. [In particular, the finiteness of the expansion (2) is justified on the basis of the boundedness of the support of the volume, i.e., the fact that f is zero-valued everywhere outside a bounded region of space. Under this assumption, we need to consider only those finitely many points $\mathbf{g}_j$ of the BCC grid for which the support of b_j has points in common with the support of f—see (2). However, a 2-D crystal that is not identically zero cannot possibly have a bounded support—see (1). Hence something else needs to be done to turn the 2-D crystal reconstruction problem into a finite system of approximate equalities such as (3) that can be solved as in (4).] For a detailed description of blobs and of their placement in space, see [21]–[23], [28]–[30]. We only note here that the recommended choices of the blobs and their placements are independent of the object to be reconstructed, rather they are based on general principles such as sampling efficiency and the desirability of being able to represent constant-valued functions. In particular, in our current work these choices are made independently of the crystal vectors of the 2-D crystal to be reconstructed.

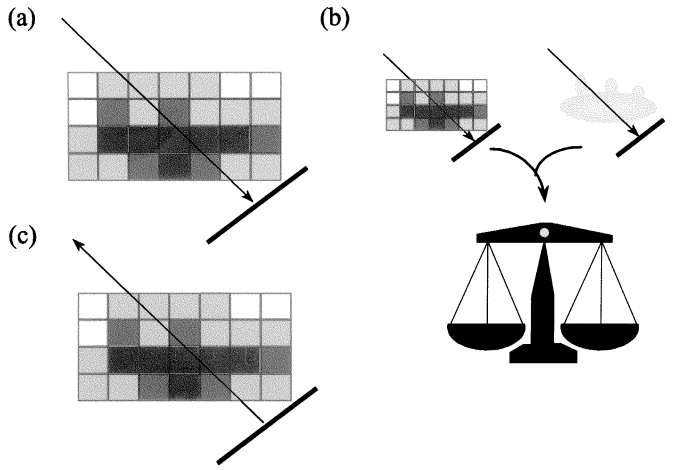


Fig. 1. Each iteration of ART consists of three steps: (a) projection of the current trial volume; (b) comparison of this projection with the experimental data; and (c) updating of the trial volume (backprojection of the difference).

F. Methodology for Objective Comparison of Algorithms

To compare the performance of the new algorithm with the traditional one, we followed the statistical methodology for objective comparison proposed in [31] and applied to electron microscopy in [23]. In this methodology, the first step is to define a task which is related to the type of information we intend to extract from the reconstructed volume. Then, several realizations from a statistically defined set of phantoms are created; these phantoms must resemble, in some way, real

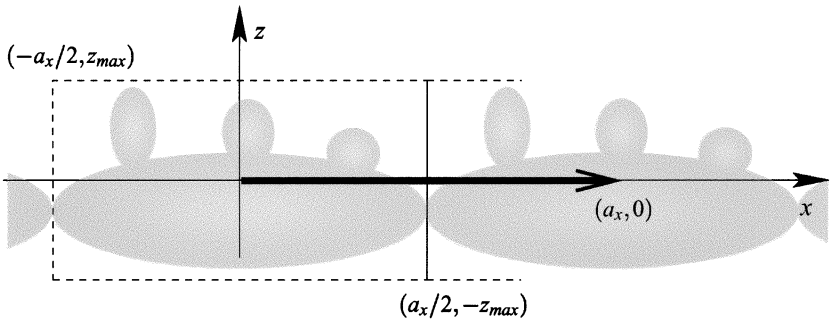


Fig. 2. A 1-D crystal: the rectangle encloses the reference unit cell. The rectangle is half-open, points on the edge of the rectangle belong to the unit cell only for the edge that is drawn solid.

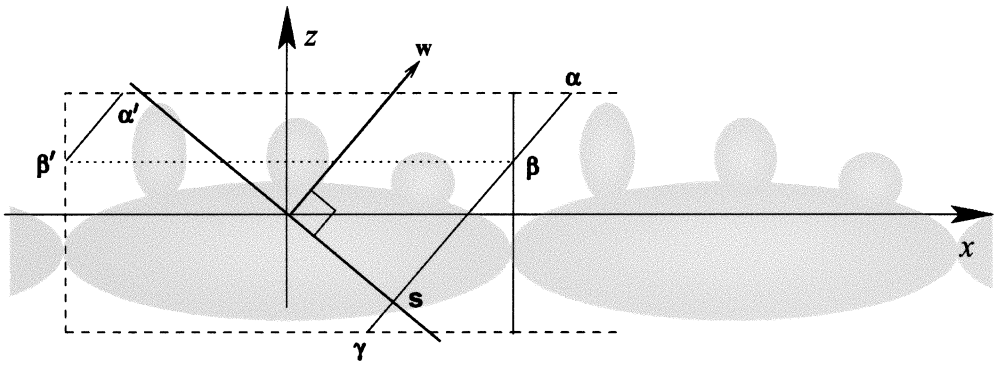


Fig. 3. According to its definition (7), $p_{\mathbf{w}}(\mathbf{s})$ is the line integral of f along the line segment from $\boldsymbol{\alpha}$ to $\boldsymbol{\gamma}$. Due to (5), the line integral of f along the line segment from $\boldsymbol{\alpha}$ to $\boldsymbol{\beta}$ is the same as the line integral of f along the line segment from $\boldsymbol{\alpha}'$ to $\boldsymbol{\beta}'$. Hence, $p_{\mathbf{w}}(\mathbf{s})$ can be calculated as the integral over the line through the reference unit cell (the union of the segments from $\boldsymbol{\alpha}'$ to $\boldsymbol{\beta}'$ and from $\boldsymbol{\beta}$ to γ).

objects of interest. After that, projections of these phantoms are generated and volumes are reconstructed from these projections using the algorithms. Finally, to measure how well the task is being performed by each algorithm, a *figure of merit* (*FOM*) is defined; this is a numerical observer measuring the degree of accomplishment of the task. Section V describes in detail the particular phantoms, projections and FOM used for this work.

III. ART for Crystals

A. Illustration on One-Dimensional Crystals

In this subsection, we present the new reconstruction algorithm applied to 1-D crystals. We do this, since the 1-D crystal reconstruction problem is sufficient for demonstrating all the essential concepts of our approach, but it is less complex mathematically than the 2-D crystal problem, and it is easier to draw illustrations of it.

A *1-D crystal* is a function f of two real variables for which exists a vector $\mathbf{a} = (a_x, 0)$, called the *crystal vector*, such that

$$f(x + ma_x, z) = f(x, z) \tag{5}$$

for all the points (x, z) and all integers m. We assume that the crystal is of a limited size in the z direction; i.e., that there is a positive number $z_{\max}$, such that

$$f(x, z) = 0, \quad \text{if } |z| \geq z_{\max}. \tag{6}$$

The set of points (ua_x, z) such that $-1/2 < u \leq 1/2$ and $-z_{\max} < z < z_{\max}$ is the *reference unit cell* of this 1-D crystal (see Fig. 2). Clearly, if we know the values of $f(x, z)$ for all points (x, z) in the reference unit cell, then we can determine $f(x, z)$ anywhere using (5) and (6).

Let $\mathbf{w}$ be the 2-D unit vector $\mathbf{w} = (w_x, w_z)$ such that $w_z \neq 0$ (see Fig. 3) and let $\mathbf{w}_\perp = (w_z, -w_x)$. The projection of a 1-D crystal f in the direction $\mathbf{w}$ is a function (that we call $p_{\mathbf{w}}$) whose arguments are the points $\mathbf{s}$ of the form $\xi\mathbf{w}_\perp$ (these lie on the line that is perpendicular to $\mathbf{w}$ and goes through the origin) and whose values are the line integrals of f along a line parallel to $\mathbf{w}$ going through $\mathbf{s}$ (again, see Fig. 3). To be exact

$$p_{\mathbf{w}}(\mathbf{s}) = \int_{-\infty}^{\infty} f(\mathbf{s} + l\mathbf{w}) dl. \tag{7}$$

Define $\mathbf{a}_{\mathbf{w}} = \mathbf{a} - \langle \mathbf{a}, \mathbf{w} \rangle \mathbf{w}$. It is easy to show that

$$p_{\mathbf{w}}(\mathbf{s} + m\mathbf{a}_{\mathbf{w}}) = p_{\mathbf{w}}(\mathbf{s}) \tag{8}$$

for all points $\mathbf{s}$ and integers m; and so $p_{\mathbf{w}}$ is also a crystal with *projection crystal vector* is $\mathbf{a}_{\mathbf{w}}$. The *reference projection unit cell* is formed by the set of points of the form $u\mathbf{a}_{\mathbf{w}}$, such that $-1/2 < u \leq 1/2$; if we know the values of $p_{\mathbf{w}}(\mathbf{s})$ for all points $\mathbf{s}$ in the reference projection unit cell, then the value of the projection at any point can be obtained by using (8). The following is an essential observation: for any $\mathbf{s}$ in the reference projection unit cell, $p_{\mathbf{w}}(\mathbf{s})$ can be calculated as the sum of line integrals along line segments within the reference unit cell of the 1-D crystal (see Fig. 3). We refer to the union of these segments as a *line through the reference unit cell*.

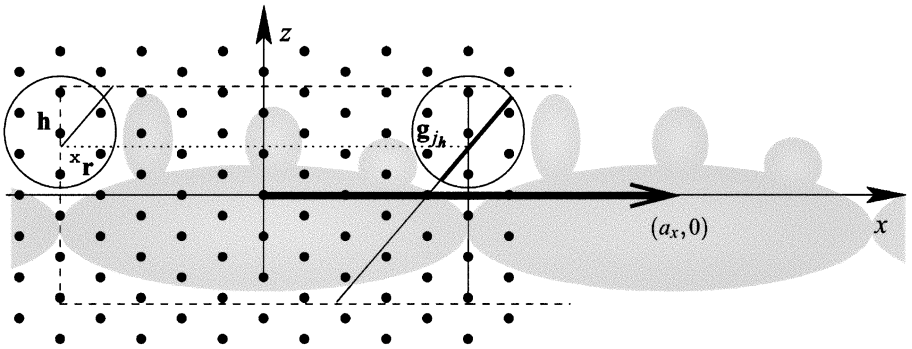


Fig. 4. Black dots and big circles represent grid points and basis functions (2-D blobs), respectively. Although the point $\mathbf{h}$ is not in G (see text), the basis functions centered at it contributes to the density of the reference unit cell. This contribution is determined by the coefficients $c_{j_\mathbf{h}}$ associated with the grid point $\mathbf{g}_{j_\mathbf{h}}$ in G such that $\mathbf{h} = \mathbf{g}_{j_\mathbf{h}} + m\mathbf{a}$ for some integer m, see (9). If the indicated line through the reference unit cell corresponds to the ith measurement, then the $l_{i,j_\mathbf{h}}$ for (3) is defined as the integral of $b_{j_\mathbf{h}}$ along the heavy line segment through it.

Suppose that, for now, the grid is *consistent* with the crystal in the sense that the grid points are arranged so that if $\mathbf{g}$ is a grid point, then so is $\mathbf{g} + m\mathbf{a}$ for any integer m. Consider the set of grid points $\mathbf{g} = (g_x, g_z)$ which satisfy both of the two criteria: (i) $-1/(2a_x) < g_x \leq 1/(2a_x)$ and (ii) the support of the basis functions centered at $\mathbf{g}$ intersects the reference unit cell; see Fig. 4. This set G is finite, and so we can write $G = \{\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_J\}$. Associating with each $\mathbf{g}_j$ a coefficient c_j, leads to a 1-D crystal (with crystal vector $\mathbf{a}$) defined as follows. Consider any point $\mathbf{r} = (x, z)$ such that $-z_{\max} < z < z_{\max}$. Let $G_\mathbf{r}$ be the set of all (finitely many) grid points $\mathbf{g}$ such that the basis function centered at $\mathbf{g}$ contains $\mathbf{r}$. For each $\mathbf{h}$ in $G_\mathbf{r}$ there is at most one $\mathbf{g}_{j_\mathbf{h}}$ in G such that $\mathbf{g} - \mathbf{g}_{j_\mathbf{h}}$ is of the form $m\mathbf{a}$ for some integer m. We define

$$f(\mathbf{r}) = \sum_{h \in G_\mathbf{r}} c_{j_\mathbf{h}} b(\mathbf{r} - \mathbf{h}) \tag{9}$$

where $c_{j_\mathbf{h}}$ is defined to be zero if there is no $\mathbf{g}_{j_\mathbf{h}}$ in G such that $\mathbf{g} - \mathbf{g}_{j_\mathbf{h}}$ is of the form $m\mathbf{a}$ for some integer m. It is easy to check that this f is indeed a 1-D crystal (with crystal vector $\mathbf{a}$) and it is completely described by the J coefficients $c_1, c_2, \ldots, c_J$. It is our intent to use such a finite description to approximate an arbitrary 1-D crystal (see Fig. 4). The coefficients are calculated based on a system such as (3), in which each y_i is a measurement within a reference projection unit cell and the $l_{i,j}$ are defined as the appropriate line integrals through the basis functions b_j.

For 1-D crystals, it is easy to select a grid that has desirable properties from the reconstruction point of view and which is also consistent with the crystal: all we need to do is to appropriately adjust the grid sampling distance. This does not carry over to 2-D crystals: since there are two crystal vectors $\mathbf{a}$ and $\mathbf{b}$, it is generally not possible to pick a sampling distance for the BCC grid so that the resulting grid is consistent with the 2-D crystal in the sense that whenever $\mathbf{g}$ is a grid point, then so is $\mathbf{g} + m\mathbf{a} + n\mathbf{b}$ for all integers m and n. However, there is an alternative way of handling lack of consistency of a grid with a 1-D crystal, and this way does carry over to 2-D crystals.

We assume that the given grid is *regular* in the sense that there is a *grid vector* $\mathbf{p} = (p_x, 0)$ such that whenever $\mathbf{g}$ is a grid point, then so is $\mathbf{g} + s\mathbf{p}$ for any integer s. Now select a real number d such that a_x/d is a nonzero integer multiple of p_x. (Although this is not needed for the theoretical discussion that follows, in practice d should be selected to be as near to 1 as possible.)

Let ${}_Df$ be defined by

$${}_Df(\mathbf{r}) = f(\mathbf{r}D) \tag{10}$$

for all points $\mathbf{r}$, where $D = \begin{pmatrix} d & 0 \\ 0 & 1 \end{pmatrix}$. The function ${}_Df$ is sometimes called a *squashed version* of f, see [28] and [32]. It is immediate from (10) that ${}_Df$ is a 1-D crystal with crystal vector $\mathbf{a}D^{-1} = (a_x/d, 0)$. It follows that the given regular grid is consistent with ${}_Df$. Now we show that all the projections of ${}_Df$ can be obtained from the projections of f. Hence we can reconstruct ${}_Df$ from these derived projections using the given grid (as discussed above for the consistent case) and then recover f using (10).

According to (7) and (10), the projection ${}_Dp_{\mathbf{w}'}$ of ${}_Df$ in direction $\mathbf{w}'$ is defined by

$${}_Dp_{\mathbf{w}'}(\mathbf{s}') = \int_{-\infty}^{\infty} f(\mathbf{s}'D + l\mathbf{w}'D)\, dl \tag{11}$$

where $\mathbf{s}'$ is a point on the line through the origin that is perpendicular to $\mathbf{w}'$. We claim that, for every 2-D unit vector $\mathbf{w}$, the unit vector

$$\mathbf{w}' = \frac{\mathbf{w}D^{-1}}{\|\mathbf{w}D^{-1}\|} \tag{12}$$

(we use $\|\cdot\|$ to denote the length of a vector) has the property that, for all points $\mathbf{s}$ on the line that goes through the origin perpendicular to $\mathbf{w}$, there is a point

$$\mathbf{s}' = -\left\langle \mathbf{w}', \mathbf{s}D^{-1} \right\rangle \mathbf{w}' + \mathbf{s}D^{-1} \tag{13}$$

on the line that goes through the origin perpendicular to $\mathbf{w}'$, such that

$${}_Dp_{\mathbf{w}'}(\mathbf{s}') = \left\|\mathbf{w}D^{-1}\right\| p_\mathbf{w}(\mathbf{s}). \tag{14}$$

[In less precise but more enlightening words: every value of the projection of f in the direction $\mathbf{w}$ provides us with a value of the projection of the squashed version of f in direction $\mathbf{w}'$ that depends only on $\mathbf{w}$ as specified in (12).] The validity of

(14) is easily tested by starting from (11), making the change of variables

$$l = \langle \mathbf{w}', \mathbf{s}D^{-1} \rangle + k \left\| \mathbf{w}D^{-1} \right\|, \tag{15}$$

substituting for $\mathbf{s}'$ using (13), making the obvious cancellation, substituting for $\mathbf{w}'$ using (12), and then observing (7).

Finally, we discuss how the y_i of (3) would be produced for the 1-D crystal case. What corresponds to an electron microscopic projection in this case is a collection of measured (and hence noisy) values of $p_{\mathbf{w}}(\mathbf{s})$ for many uniformly spaced points $\mathbf{s}$; only a few of these are points inside the reference projection unit cell. Similar to what is indicated in Section II-C, in order to reduce the noise, we should make use of all the measured data. This can be achieved by taking a 1-D Fourier transform of $p_{\mathbf{w}}(\xi \mathbf{w}_{\perp})$ with respect to the variable ξ. This results, after some 1-D processing, in values only at points whose spacing is $\|\mathbf{a}_{\mathbf{w}}\|^{-1}$; that is, the inverse of the length of the reference projection unit cell. Only the central $2M+1$ values, for some positive integer M, are reliable (due to the sampled and noisy nature of the measured data), and using the inverse discrete Fourier transform only on these values provides us with estimates of the $p_{\mathbf{w}}(\mathbf{s})$ at $2M+1$ equally spaced points within the reference projection unit cell. The accuracy of these estimates is much better than that of the raw measurements, due to the fact that they are calculated based on all the measurements (and not just those taken within the reference projection unit cell).

B. ART for 2-D Crystals

In this subsection the above-presented 1-D crystal reconstruction algorithm is generalized to the more relevant 2-D crystal case. The development in this subsection follows closely that in the previous subsection, except that we now assume that f is a 2-D crystal with crystal vectors $\mathbf{a}$ and $\mathbf{b}$.

Let $\mathbf{w} = (w_x, w_y, w_z)$ be a 3-D unit vector such that $w_z \neq 0$. The projection $p_{\mathbf{w}}$ of f in the direction $\mathbf{w}$ is defined by the formula given in (7), but now the interpretation of $\mathbf{s}$ is that it is a point in the plane through the origin that is perpendicular to $\mathbf{w}$. It is the case that $p_{\mathbf{w}}$ is a crystal with crystal vectors

$$\begin{aligned} \mathbf{a}_{\mathbf{w}} &= \mathbf{a} - \langle \mathbf{a}, \mathbf{w} \rangle \mathbf{w} \\ \mathbf{b}_{\mathbf{w}} &= \mathbf{b} - \langle \mathbf{b}, \mathbf{w} \rangle \mathbf{w}. \end{aligned} \tag{16}$$

This can be easily proven by putting $\mathbf{s} + m\mathbf{a}_{\mathbf{w}} + n\mathbf{b}_{\mathbf{w}}$ in place of $\mathbf{s}$ in (7) (with integers m and n) and changing the variable l to $k = l - m\langle \mathbf{a}, \mathbf{w} \rangle - n\langle \mathbf{b}, \mathbf{w} \rangle$. The *reference projection unit cell* is formed by the set of points of the form $u\mathbf{a}_{\mathbf{w}} + v\mathbf{b}_{\mathbf{w}}$, such that $-1/2 < u \leq 1/2$ and $-1/2 < v \leq 1/2$. It is not difficult to see that the situation depicted in Fig. 3 generalizes to 2-D crystals and consequently, for any $\mathbf{s}$ in the reference projection unit cell, $p_{\mathbf{w}}(\mathbf{s})$ can be calculated as the sum of line integrals along line segments within the reference unit cell of the 2-D crystal (i.e., as an integral along a line through the reference unit cell).

In strict analogy to what was done before, we say that a grid (which is now a set of points in 3-D space) is *consistent* with the 2-D crystal with crystal vectors $\mathbf{a}$ and $\mathbf{b}$ if, whenever $\mathbf{g}$ is a grid point, then so is $\mathbf{g} + m\mathbf{a} + n\mathbf{b}$ (for any integers m and n). It is again not hard to see how to carry over the ideas depicted in Fig. 4 to 2-D crystals, which subsequently allows us to recover (using ART) a 2-D crystal from measurements within the reference projection unit cell. However, such a reconstruction is possible only for grids consistent with the 2-D crystal, and there is no guarantee that some "desirable" grid (such as the BCC grid) will be consistent with f.

To show how we overcome this difficulty, we now discuss how the essential idea of squashing carries over to 2-D crystals. For any invertible matrix D of the form

$$D = \begin{pmatrix} d_{11} & d_{12} & 0 \\ d_{21} & d_{22} & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{17}$$

we can define ${}_Df$ using (10), bearing in mind that $\mathbf{r}$ is now a point in 3-D space. It is trivial to prove that ${}_Df$ is a 2-D crystal with crystal vectors $\mathbf{a}D^{-1}$ and $\mathbf{b}D^{-1}$.

Suppose that we are given a *regular* grid; i.e., one that has, for two fixed linearly independent *grid vectors* $\mathbf{p} = (p_x, p_y, 0)$ and $\mathbf{q} = (q_x, q_y, 0)$, the property that whenever $\mathbf{g}$ is a grid point, then so is $\mathbf{g} + s\mathbf{p} + t\mathbf{q}$ (for any integers s and t). Since the BCC grid is obtained by interlacing two cubic grids, it is clearly regular (and we can, in fact, choose $p_y = 0$, $q_x = 0$ and $p_x = q_y$). Suppose that s_a, t_a, s_b and t_b are four integers such that the vectors $s_a\mathbf{p} + t_a\mathbf{q}$ and $s_b\mathbf{p} + t_b\mathbf{q}$ are linearly independent. Then there is a unique invertible matrix D of the form shown in (17) that satisfies

$$\begin{aligned} \mathbf{a}D^{-1} &= s_a\mathbf{p} + t_a\mathbf{q} \\ \mathbf{b}D^{-1} &= s_b\mathbf{p} + t_b\mathbf{q}. \end{aligned} \tag{18}$$

(This follows from the property of the crystal vectors that they are linearly independent.) It is now easy to show that the given regular grid is consistent with ${}_Df$. We leave for later the all-important consideration of how the s_a, t_a, s_b, and t_b should be chosen and turn instead to discussing the relationship between projections of ${}_Df$ and projections of f.

It turns out that having got our formalism for 2-D crystals to match that for 1-D crystals, we can now obtain a matching result (and using the same derivation): namely, that for the 3-D unit vector $\mathbf{w}$ there is a unit vector $\mathbf{w}'$ defined by (12) such that, for all points $\mathbf{s}$ on the plane that goes through the origin perpendicular to $\mathbf{w}$, there is a point $\mathbf{s}'$ defined by (13) on the plane that goes through the origin perpendicular to $\mathbf{w}'$ such that (14) is satisfied. So from the actual projections of f, we can obtain projections of ${}_Df$, from which we can reconstruct ${}_Df$ using the given regular grid for the centers of our blobs, and then we can calculate f using (10).

We complete this subsection by discussing how the y_i of (3) are produced for the 2-D crystal case. The first phase of this process (that is shared by the traditional approach and our approach) has been summarized in Sections II-C and II-D. At the end of that phase the MRC suite of programs outputs estimates of the crystal vectors $\mathbf{a}$ and $\mathbf{b}$ and, for each projection, an estimate of the associated direction vector $\mathbf{w}$ (that, by (16), also yields $\mathbf{a}_{\mathbf{w}}$ and $\mathbf{b}_{\mathbf{w}}$) and a 2-D array of complex numbers $b_{K,L}$ ($-M \leq K \leq M, -N \leq L \leq N$), which are the estimated

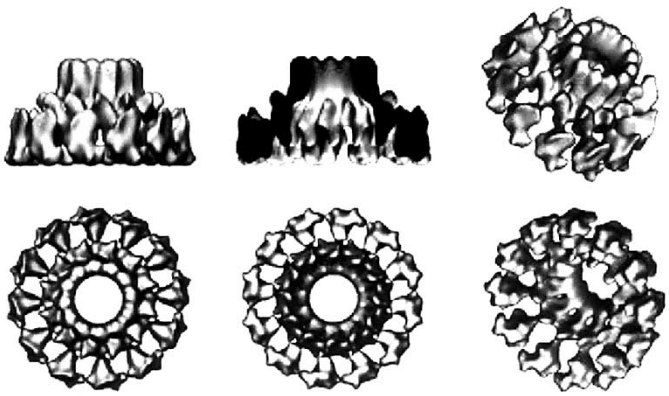


Fig. 5. Several views of a 3-D surface rendering of the bacteriophage $\phi 29$ connector, based on a reconstruction from real data [36].

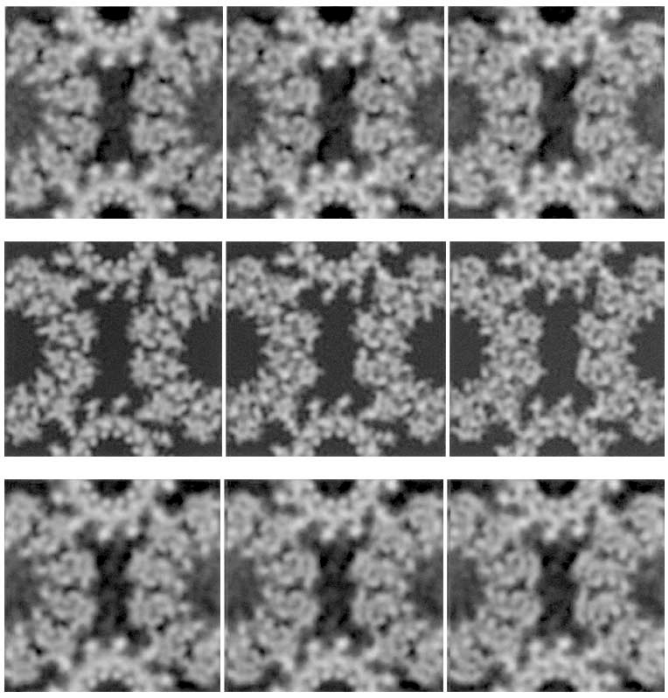


Fig. 6. Reconstructions of a phantom of the bacteriophage $\phi 29$ connector using the traditional algorithm (top) and the new algorithm (bottom), and low-pass filtered version of a density volume of the same structure created from atomic coordinates (middle). In each row three corresponding consecutive slices are shown.

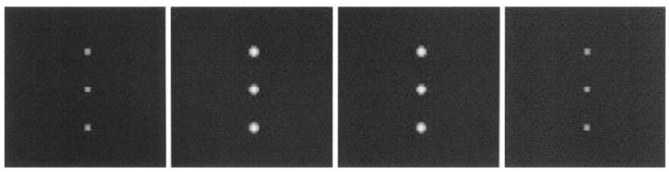


Fig. 7. The four central slices of a $64 \times 64 \times 64$ digitization of a phantom composed of three small spheres ($\mathrm{radius} = 1.75$ times the edge of a voxel). The density of a voxel is proportional to the volume of the voxel's intersection with the small spheres (being equal to one when it is totally inside one of them). The image shows planes perpendicular to the y-axis (the tilt axis is parallel to the y-axis).

values of the spots in the FT of the projection. A more precise (and detailed) description of the nature of the $b_{K,L}$ is given in [33, Sec. V-A]. In the same place (see [33, eq. (14)]), it is stated

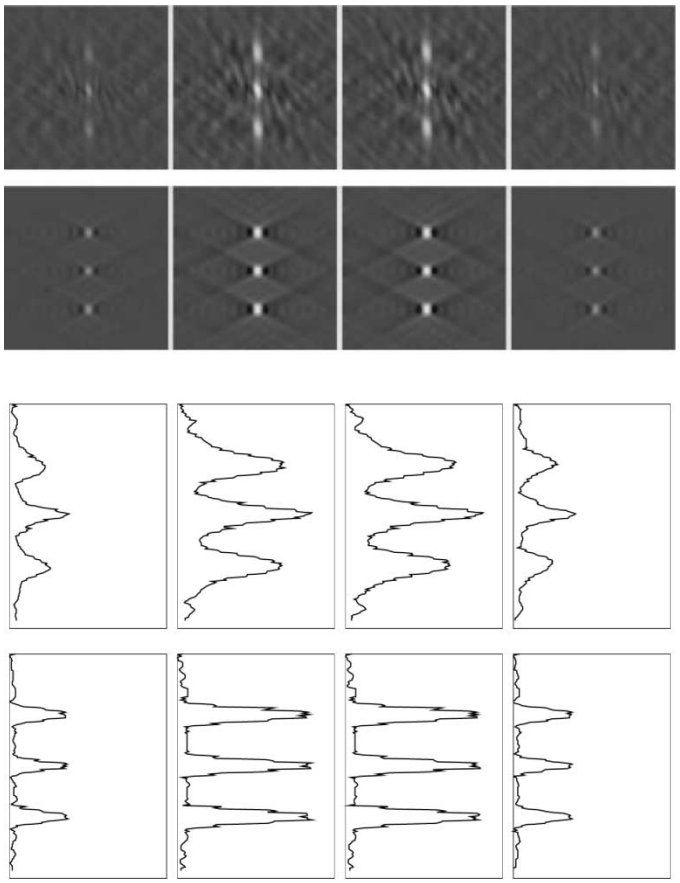


Fig. 8. The two upper rows are central slices (matching those in Fig. 7) of 3-D reconstructions obtained using the MRC package. They were obtained from 13 (first row) and 49 (second row) noiseless projections, respectively. The gamma of the images was adjusted so that differences in the background are more noticeable. The two lower rows are plots matching the central column for each slice. The third row plots correspond to the first row of images and the fourth row plots to the second row of images.

that $p_{\mathbf{w}}(\mathbf{s})$ can be estimated for certain points $\mathbf{s}$ from the $b_{K,L}$ by the use of the inverse discrete FT; namely, for $-M \leq k \leq M$ and $-N \leq l \leq N$,

$$p_{\mathbf{w}}\left(\frac{k}{2M+1}\mathbf{a}_{\mathbf{w}} + \frac{l}{2N+1}\mathbf{b}_{\mathbf{w}}\right) \simeq \sum_{K=-M}^{M}\sum_{L=-N}^{N} b_{KL} e^{2\pi\sqrt{-1}((Kk/2M+1)+(Ll/2N+1))}. \quad (19)$$

This provides us with high-quality estimates of the values of the projection at uniformly sampled points within the reference projection unit cell.

IV. Practical Considerations

A. Accuracy and Efficiency of the Implementation

At first sight, it may appear that the process described in the last section may necessitate numerous interpolations that may not only be demanding on computational resources but, worse, may also be a source of inaccuracies due to interpolation errors. In this subsection we show that this is in fact not the case: the process involving squashing can be implemented without any additional interpolations (as compared to the process that would

be used if the given grid happened to be consistent with the 2-D crystal).

Having obtained the projection data at the points $\mathbf{s}$ indicted in (19), we need to calculate ${}_D p_{\mathbf{w}'}(\mathbf{s}')$ for points $\mathbf{s}'$ in the reference projection unit cell of the projection of ${}_D f$ in the direction $\mathbf{w}'$. By (16), the crystal vectors of the projection of ${}_D f$ in the direction $\mathbf{w}'$ are ${}_D\mathbf{a}_{\mathbf{w}'} = \mathbf{a}D^{-1} - \langle \mathbf{a}D^{-1}, \mathbf{w}' \rangle \mathbf{w}'$ and ${}_D\mathbf{b}_{\mathbf{w}'} = \mathbf{b}D^{-1} - \langle \mathbf{b}D^{-1}, \mathbf{w}' \rangle \mathbf{w}'$. By applications of (12)–(14), one can easily show that, for $-M \leq k \leq M$ and $-N \leq l \leq N$

$$
\begin{aligned}
{}_D p_{\mathbf{w}'} &\left(\frac{k}{2M+1}\, {}_D\mathbf{a}_{\mathbf{w}'} + \frac{l}{2N+1}\, {}_D\mathbf{b}_{\mathbf{w}'} \right) \\
&= \left\| \mathbf{w}D^{-1} \right\| p_{\mathbf{w}} \left(\frac{k}{2M+1}\mathbf{a}_{\mathbf{w}} + \frac{l}{2N+1}\mathbf{b}_{\mathbf{w}} \right); \quad (20)
\end{aligned}
$$

and so from the left-hand side of (19), we immediately obtain the values ${}_D p_{\mathbf{w}'}$ at uniformly sampled points within its reference unit cell. No interpolations are needed, only a multiplication by a value that is constant for each projection (but changes from one projection to another).

The next step in the process is the reconstruction, using ART with blobs as the basis functions, of ${}_D f$. At the end of this, we obtain a J-dimensional vector $\mathbf{c}$ that gives rise to an approximation of ${}_D f$ of the form (see (2))

$$
{}_D f(\mathbf{r}) \approx \sum_{j=1}^{J} c_j b(\mathbf{r} - \mathbf{g}_j). \quad (21)
$$

Using (10), this yields

$$
f(\mathbf{r}) \approx \sum_{j=1}^{J} c_j\, {}_{D^{-1}} b(\mathbf{r} - \mathbf{g}_j D) \quad (22)
$$

in which the *desquashed blob* ${}_{D^{-1}} b$ is defined in the manner of (10). Thus, in order to get from the squashed reconstruction to the actual reconstruction, we do not need to recalculate the coefficients of the expansion into basis function but rather we replace the basic blob by its desquashed version and change the locations of the grid points. In practice, ${}_D f$ is not explicitly produced, we use only the vector $\mathbf{c}$ defining it by (21) to obtain f directly from (22).

Typically, given a vector $\mathbf{c}$, the function f represented by (2) is evaluated for points $\mathbf{r}$ placed on a cubic grid, by precalculating the values of b over its support for a much finer cubic grid and then calculating the contribution of b_j to f by adding values obtained from a regularly subsampled version of this finer array. Hence, as long as the support of ${}_{D^{-1}} b$ is not essentially larger than that of b, the evaluation of (22) is no more expensive (or less accurate) than the evaluation of (2).

B. Choice of the Squashing Matrix

In the previous section, we did not specify exactly how the s_a, t_a, s_b, and t_b, and hence the resulting squashing matrix D of (17), should be chosen. We now return to this important point.

As can be seen from the previous subsection, especially (22), D distorts both the basis functions and the grid. Since presumably we start with basis functions and a grid that are considered "good" based on previous experience [21], [29], it seems desirable to keep these distortions to be as small as possible; in other words to make D to be as nearly the identity matrix as possible subject to satisfying (18) for some appropriate integers s_a, t_a, s_b and t_b. (Rotation matrices are equally acceptable, but we will not get into that discussion here.) This leads to the following recommendation: subject to the condition that the vectors $s_a\mathbf{p} + t_a\mathbf{q}$ and $s_b\mathbf{p} + t_b\mathbf{q}$ should be linearly independent, choose the integers s_a, t_a, s_b and t_b so that $\|\mathbf{a} - s_a\mathbf{p} - t_a\mathbf{q}\|^2 + \|\mathbf{b} - s_b\mathbf{p} - t_b\mathbf{q}\|^2$ is as small as possible. (With this choice of D, the end points of the crystal vectors of ${}_D f$ will be those grid points that are nearest to the end points of the crystal vectors of f.)

C. Electron Microscopic Issues

In this paper, we concentrate on the image processing aspects of our approach and largely ignore the practical problems that arise from the nature of the data that is obtained by TEM of a biological macromolecule. In this subsection, we briefly mention two relevant considerations but leave it to the reader to find detailed answers in the technical literature on electron microscopy.

The motif of a 2-D crystal may have various symmetries and one can make use of these to improve the quality of the reconstructions; for example, by using as input not only the projection data recorded by the electron microscope but also those projections that can be generated by symmetry. The reconstruction algorithm needs to know the direction of and values in these extra projections. When they are generated by rotation or reflection it is trivial to obtain all the projection directions (just applying the symmetry to $\mathbf{w}$) and the values (there are no changes to them). The case of a *glide reflection* (a reflection followed by a shift) is slightly more complex, but can still be handled within our framework. (Glide reflections have been tabulated for all crystallographic symmetries—see [34].) In some cases we found it convenient for the handling of symmetries in conjunction with the BCC grid to insist that the t_a and the s_b in the last subsection should both be chosen to be zero.

Another technical concern is the relationship between the coordinate system assumed by our mathematical discussion and the coordinate system attached to the electron micrographs. Even if we know the projection direction $\mathbf{w}$, it is not trivial to identify the point in an electron micrograph that corresponds to a particular $\mathbf{s}$ in (7), since typically we have no prior knowledge regarding the in-plane rotation of the electron micrograph. This mandates that careful attention has to be paid to overcome any discrepancy between the experimental coordinate system and that assumed by (19).[3]

V. Results

A. Biological Motivation

Prior to describing and applying our methodology for systematic comparison of the reconstruction algorithms, we show the results of reconstructing a very realistic phantom. This phantom was produced from the structure of a protein (the connector of

[3]Details of how such technical problems are handled can be found in the source code for our algorithm that is available as a part of the Xmipp package [35]. This is available at http://www.cnb.uam.esuam.es/~bioinfo.

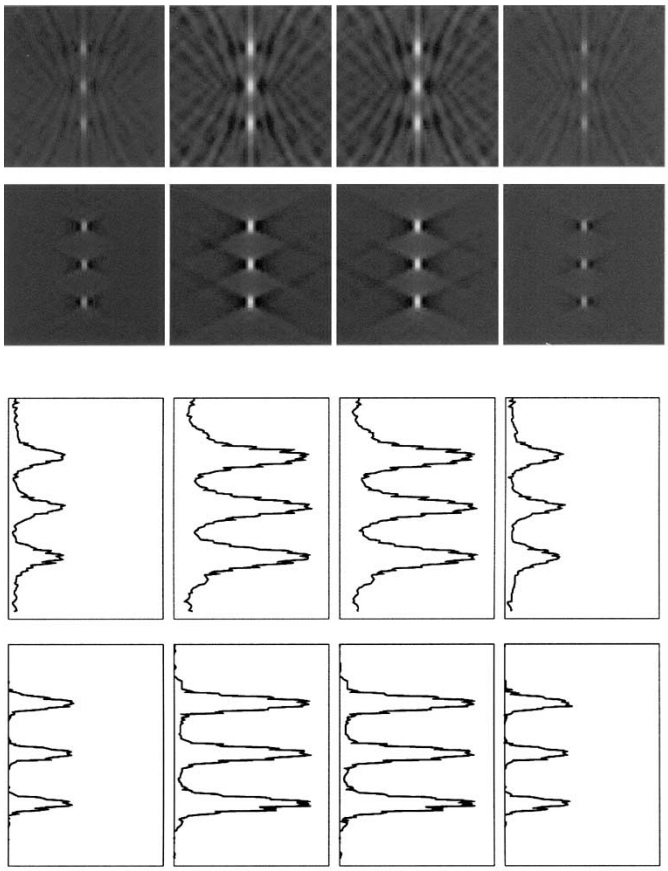


Fig. 9. The two upper rows are central slices (matching those in Fig. 7) of 3-D reconstructions obtained using the new ART with blobs algorithm. They were obtained from 13 (first row) and 49 (second row) noiseless projections, respectively. The gamma of the image was adjusted in the same way as in Fig. 8. The two lower rows are plots matching the central column for each slice. The third row plots correspond to the first row of images and the fourth row plots to the second row of images.

the bacteriophage $\phi 29$) that is known at atomic resolution. A 3-D reconstruction of this structure obtained from electron microscopy data [36] is shown in Fig. 5.

In our model the atoms are replaced by small spheres.[4] The number of projections, the projection angles and the MRC processing are identical to what is described in [36] for real data. The reconstructions are shown in Fig. 6. We show corresponding slices of the reconstruction by the MRC package, of the reconstruction by new algorithm, and of a low pass filtered version of the phantom. The results obtained by the two algorithms for this test case are quite similar, although not identical. The obvious question is: Which algorithm provides a more accurate description? The remainder of this section is devoted to answering this question; we present various experiments that we performed to compare the *standard* Fourier space method (based on the MRC approach) and the new ART with blobs algorithm.

B. Comparison of the 3-D Reconstruction Algorithms

1) Phantom Generation: The first set of experiments uses a phantom composed of little spheres (see Fig. 7). These experiments are mainly exploratory, and were designed for the purpose of finding out how many projections are needed to properly sample the Fourier space. We were also interested in how a trivial structure (practically a point) is reconstructed, since algorithm-induced artifacts are more clearly identifiable in this kind of simple phantoms. Such information is helpful in designing the experiments for a more thorough algorithm comparison.

The second set of experiments uses phantoms composed of cylinders and spheres, and addresses the topic of detectability of small structures. (An example of such a phantom can be seen in Fig. 12.)

The phantoms are mathematically defined in the sense that the file storing them contains the crystal vectors, the crystal size (the number of unit cells), and the parameters that define each feature (either a cylinder or a sphere) contained in the motif.

2) Projection Generation: The simulated data-collection geometry was single axis tilting [38], with maximum tilt angle $\pm 60^\circ$. (To be more precise, for all projections in our experiments, the w_y of Section III-B is equal to 0.) Special care was taken to generate realistic noise by considering the addition of noise not only to the pixel values of the projections, but also to the distances between neighboring unit cells (*lattice distortion*). We also incorporated *feature skipping,* meaning that for each unit cell of the phantom and for each feature (cylinder or sphere) belonging to this unit cell a random number was generated, and that feature was skipped (not taken into account when computing the projection) if that number was less than some fixed threshold.

As discussed in the last paragraph of Section III-B, at the end of the first phase of processing the MRC suite of programs [15] (in our experiments also helped by SPECTRA [18]) outputs estimates of the crystal vectors $\mathbf{a}$ and $\mathbf{b}$ and, for each projection, an estimate of the associated direction vector $\mathbf{w}$ and a 2-D array of complex numbers $b_{K,L}$, which are the estimated values of the spots in the FT of the projection. In our experiments this output was generated by two alternative approaches: one generating noiseless data and the other noisy data. For noiseless data, the $b_{K,L}$ were directly calculated from the analytic 3-D FT of the phantom. For noisy data, real space noisy projections were created and processed using the MRC and SPECTRA packages to provide estimated values of the $\mathbf{a}$, $\mathbf{b}$, $\mathbf{w}$ and $b_{K,L}$.

3) Results With the Phantom of Spheres:

Noiseless Data: The phantom described in Fig. 7 was reconstructed from two different sets of noiseless projections using the traditional (MRC) algorithms based on interpolation in Fourier space (Fig. 8) and also using our new approach (Fig. 9). Since the conclusions obtained from examining the results for this noiseless case are identical those for the noisy case we postpone their discussion until the noisy case is presented.

a) Noisy Data: Similar experiments were done after adding noise with mean zero and standard deviation 10 to the projection pixel values, using mean zero and standard deviation half of a voxel edge for the lattice distortion, and 3% for the feature skip. Three reconstructions were performed with each algorithm using 13, 49, and 97 projections, respectively—see Figs. 10 and 11. (We mention that 90 projections are needed to satisfy the criterion of [19], assuming that the thickness of the specimen is the thickness of the reconstructed volume.)

[4]The atomic coordinates can be found [Online] at: http://www.pdb.org under the PDB ID: 1FOU [37].

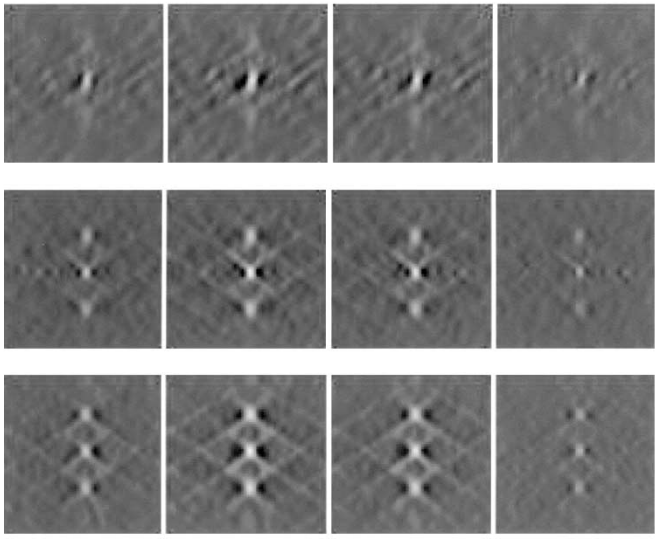


Fig. 10. Central slices (matching those in Fig. 7) of 3-D reconstructions obtained using the MRC package. The reconstruction were obtained from 13 (top), 49 (middle) and 97 (bottom) noisy projections. The gamma of the image was adjusted in the same way as in Fig. 8.

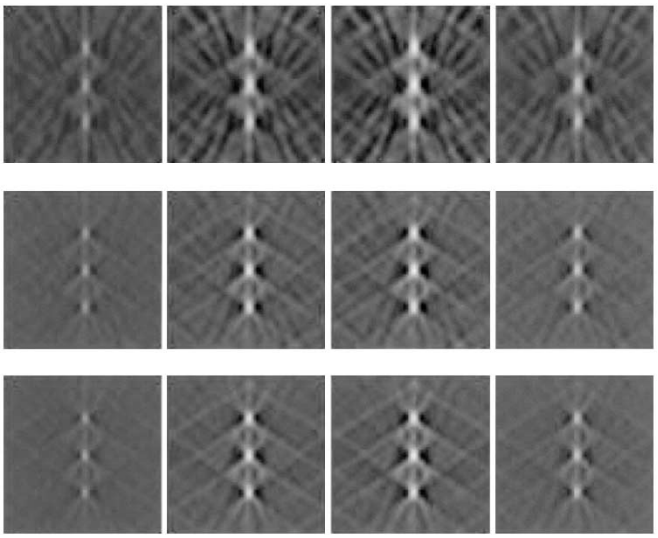


Fig. 11. Central slices (matching those in Fig. 7) of 3-D reconstructions obtained using the new ART with blobs algorithm. The reconstructions were obtained from 13 (top), 49 (middle) and 97 (bottom) noisy projections. The gamma of the image was adjusted in the same way as in Fig. 8.

The reconstructions suggest that the accuracy of the traditional algorithm (at least for this particular phantom) depends on the position of the feature when the number of projections used is small (the central sphere in the reconstruction is better defined than the spheres above and below it), while this is not the case for ART with blobs.

The following is our explanation of this experimentally observed behavior of the algorithms. The magnitude of the FT of a small sphere varies slowly. If the sphere is centered at the origin, then the phase is also slowly varying and the FT is easy to interpolate. However, the further the sphere is located from the center, the more rapidly will the phase of its FT change and the more difficult it will be to estimate the FT by interpolation. On the other hand, ART solves essentially the same system of equations regardless of the location of the spheres.

Fig. 12. Computer generated rendering of a $64 \times 64 \times 64$ voxel digitization of a phantom consisting of a hollow cylinder (inner radius $= 20$ voxel edges, outer radius $= 29$ voxel edges) and 24 little spheres (radius $= 1.50$ voxel edges). The spheres are located around the exterior of the cylinder at each 15 degrees, the value of their z coordinate is ± 4.25 voxel edges and, for a given (x, y) value, only one sphere is allowed (that is, either there is a sphere at the top of the cylinder or at the bottom). Consequently, there are 2^{24} possible phantoms.

TABLE I
STANDARD DEVIATIONS (TOP) AND AVERAGES (MIDDLE) OF THE FOM FOR THE DIFFERENT VALUES OF THE PARAMETERS FOR ART-BASED RECONSTRUCTIONS. THE LABELS IN BOLD DENOTE THE VALUES OF λ (HORIZONTAL) AND NUMBER OF CYCLES THROUGH THE DATA (VERTICAL). THE MAXIMUM VALUE OF THE FOM AND THE CORRESPONDING VARIANCE ARE UNDERLINED. IN THE PLOT OF THE AVERAGES (BOTTOM) THE AXES ARE NOT TO SCALE

	0.0004	**0.002**	**0.004**	**0.02**	**0.04**	**0.1**
1	1.17	1.20	1.17	1.24	1.97	1.83
4	1.17	1.13	1.12	1.13	1.10	1.84
8	1.17	1.10	1.17	**1.04**	1.36	1.86
16	1.13	1.19	1.02	1.18	1.55	1.97
24	1.12	1.01	1.03	1.23	1.63	2.52
32	1.10	1.03	1.03	1.30	1.79	1.93

	0.0004	**0.002**	**0.004**	**0.02**	**0.04**	**0.1**
1	21.54	21.50	21.54	21.71	21.64	21.79
4	21.57	21.64	21.82	22.39	22.36	21.87
8	21.61	21.79	22.04	**22.75**	22.00	21.96
16	21.61	22.00	22.32	22.29	21.61	21.64
24	21.68	22.14	22.57	21.86	21.18	20.50
32	21.79	22.21	22.57	21.72	20.21	20.50

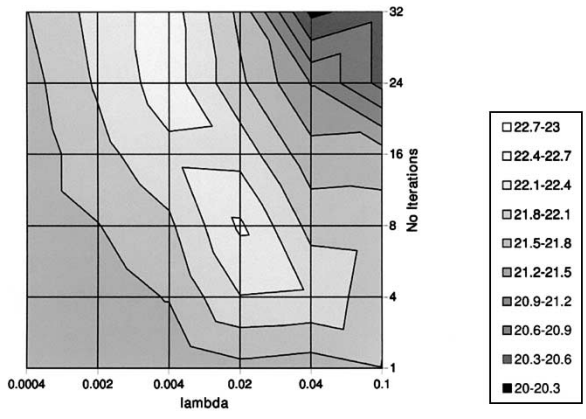
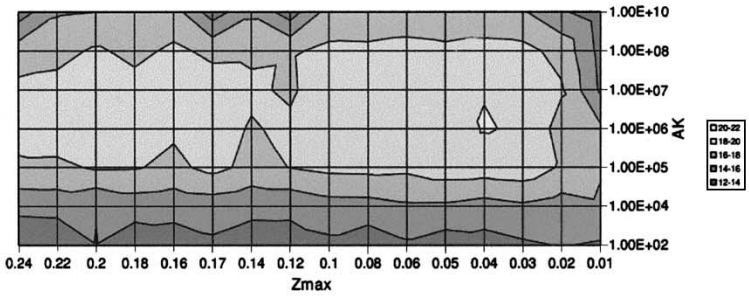


TABLE II
STANDARD DEVIATIONS (TOP) AND AVERAGES (MIDDLE) OF THE FOM FOR THE DIFFERENT VALUES OF THE PARAMETERS FOR MRC-BASED RECONSTRUCTIONS. THE LABELS IN BOLD DENOTE THE VALUES OF AK (VERTICAL) AND ZMAX (HORIZONTAL). THE MAXIMUM VALUE OF THE FOM AND THE CORRESPONDING VARIANCE ARE UNDERLINED. IN THE PLOT OF THE AVERAGES (BOTTOM) THE AXES ARE NOT TO SCALE

	0.24	**0.22**	**0.20**	**0.18**	**0.17**	**0.16**	**0.14**	**0.12**	**0.10**	**0.08**	**0.06**	**0.05**	**0.04**	**0.03**	**0.02**	**0.01**
10^{10}	2.32	2.57	2.57	2.28	2.71	2.32	2.28	2.32	2.48	2.40	2.46	2.24	2.18	2.62	2.27	2.57
10^{8}	2.61	1.85	2.47	2.56	2.11	1.97	2.34	2.20	2.33	2.31	2.12	1.81	1.83	1.97	2.08	2.54
10^{7}	2.24	1.81	2.10	2.47	1.85	1.69	1.75	2.29	1.79	1.93	1.98	1.82	1.74	1.71	1.62	1.94
10^{6}	1.90	1.76	2.31	1.93	2.17	2.18	2.39	2.23	2.20	1.85	2.01	1.56	**1.93**	1.95	2.25	2.59
10^{5}	2.07	2.06	2.20	2.03	2.15	2.12	2.15	2.38	1.90	1.99	2.35	1.84	1.91	1.90	1.76	2.09
10^{4}	2.06	1.84	2.49	2.50	2.68	2.70	2.56	2.44	1.93	2.53	2.14	2.25	2.41	2.21	2.66	1.59
10^{2}	2.85	2.29	2.01	2.51	2.31	2.39	2.76	1.93	2.79	2.94	2.32	2.81	2.90	2.17	2.57	1.85

	0.24	**0.22**	**0.20**	**0.18**	**0.17**	**0.16**	**0.14**	**0.12**	**0.10**	**0.08**	**0.06**	**0.05**	**0.04**	**0.03**	**0.02**	**0.01**
10^{10}	12.96	15.79	16.00	16.18	16.96	12.96	15.61	12.96	16.86	16.96	17.04	16.71	16.61	16.21	14.25	12.93
10^{8}	17.07	17.43	18.25	17.54	18.32	17.64	17.68	17.79	18.36	18.32	18.54	18.39	18.68	18.57	17.36	13.36
10^{7}	18.46	18.61	18.57	18.64	18.32	18.79	18.36	17.68	19.32	19.32	18.93	19.50	19.93	18.89	18.21	13.68
10^{6}	18.71	18.93	18.64	18.54	18.29	18.89	17.82	18.43	18.89	19.32	19.21	19.07	**20.11**	19.79	17.18	16.36
10^{5}	17.68	17.64	18.18	18.14	17.54	18.54	17.46	18.04	18.50	18.57	18.43	19.07	19.04	19.25	17.75	16.23
10^{4}	14.64	14.75	14.04	14.89	14.68	14.89	14.43	14.46	15.21	15.25	15.71	15.68	15.21	15.68	15.11	15.64
10^{2}	12.11	12.29	13.96	12.71	13.07	13.68	13.25	13.21	13.57	12.68	13.71	12.93	13.21	13.71	14.11	13.79



4) Results With Cylindrical Phantoms: Another experiment was performed using the family of phantoms described in detail in Fig. 12. The number of noisy projections created in this case was 97 (satisfying the criterion of [19]). The task to be performed based on a reconstruction from the projections is to identify, as closely as possible, which of the 2^{24} possible phantoms was used to generate the projection data. To decide whether a particular sphere in a particular reconstruction is at the top or at the bottom, the average reconstructed value of all the voxels that would be inside the sphere if it were at the top is computed and compared with the average reconstructed value of all the voxels that would be inside the sphere if it were at the bottom. If the former value is bigger (smaller) than the latter, then the sphere is estimated to be at the top (bottom). The FOM used to measure the quality of the reconstruction is the number of spheres correctly assigned, therefore its maximum value is 24 and its minimum is 0.

Since each method has its own parameters that need to be tuned, the experiment is performed in two stages. In the first one, called *training*, the optimal values of the parameters of the algorithms are estimated. In the second one, called *testing*, the algorithms (each with their parameters fixed at the estimated optima) are compared. Different sets of 28 randomly selected phantoms were used for the training and the testing.

Training: The parameters to be optimized for ART (4) are the relaxation parameter λ and the number of cycles through the data. For the MRC approach (in particular, for the LATLINE program described in [19]), AK and ZMAX need to be optimized (AK is the relative weight for phases with respect to magnitudes of the projection FT, and ZMAX is the maximum z^* value expected for the data set; for a detailed description see [19]). Tables I and II summarize the results for ART and MRC.

We remark that the standard deviation of the FOMs for the different values of the parameters are greater for the fully MRC-based method. (In general, a low standard deviation is a desired property, since it implies better reproducibility.)

Testing: The average values of the FOM were 19.89 and 22.41 for MRC and ART, respectively. A *t*-test over paired samples [39] was performed with the conclusion that the hypothesis that ART performs better than MRC (from the point of view of the FOM) can be maintained up to a significance of 0.9999 ($P < 0.0001$).

VI. DISCUSSION

Our main motivation for implementing a reconstruction algorithm for crystals that works in real space and is iterative is not to replace the well-established methods used at present

but to create a framework in which we can easily incorporate information available from other sources. This information, such as surface reliefs from atomic force or shadowing microscopies (see [33] and [40] for a description of our work in this area), or a more suitable model of the imaging device (contrast transfer function—see [33]), is more easily and naturally incorporated in a real space iterative method than in a Fourier space based noniterative algorithm. For example, in the ART-based method, the contrast transfer function can be taken into account by modifying the $l_{i,j}$ in the system matrix [see (3)] without any modification of the algorithm's logic. Surface relief information translates into constraints of the type $f(\mathbf{r}) = 0$, for points $\mathbf{r}$ that are outside the surface; this naturally combines with (2) to give us additional equations in the unknown coefficients c_j, and the extended system of equations can be solved using ART.

However, our tests indicate that the new algorithm behaves better than the traditional one, even without the incorporation of such additional information, especially when the number of projections is small. As the number of projections increases, the results provided by the two algorithms become more similar but not identical. The experiment with the cylindrical phantoms suggests that even then ART performs better than the Fourier-based method (from the point of view of the FOM used), and also presents greater reliability (since the variance of the FOM is smaller, see Tables I and II).

ACKNOWLEDGMENT

Dr. R. M. Lewitt's wisdom and patience has been fundamental for the completion of this work, which has also been greatly improved by fruitful discussions with Dr. E. Dimmeler.

REFERENCES

[1] R. Henderson, J. M. Baldwin, T. A. Ceska, F. Zemlin, E. Beckmann, and K. H. Downing, "Model for the structure of bacteriorhodopsin based on high-resolution electron cryo-microscopy," *J. Mol. Biol.*, vol. 213, pp. 899–929, 1990.

[2] W. Kühlbrandt, D. N. Wang, and Y. Fujiyoshi, "Atomic model of plant light-harvesting complex by electron crystallography," *Nature*, vol. 367, pp. 614–621, 1994.

[3] E. Nogales, S. G. Wolf, and K. H. Downing, "Visualizing the secondary structure of tubulin: Three-dimensional map at 4 Å," *J. Struct. Biol.*, vol. 118, pp. 119–127, 1997.

[4] K. Mitsuoka, T. Hirai, K. Murata, A. Miyazawa, A. Kidera, Y. Kimura, and Y. Fujiyoshi, "The structure of bacteriorhodopsin at 3.0 Å resolution based on electron crystallography: Implication of the charge distribution," *J. Mol. Biol.*, vol. 286, no. 3, pp. 861–82, 1999.

[5] K. Murata, K. Mitsuoka, T. Hirai, T. Walz, P. Agre, J. Heymann, A. Engel, and Y. Fujiyoshi, "Structural determinants of water permeation through aquaporin-1," *Nature*, vol. 407, pp. 599–605, 2000.

[6] G. Ren, A. Cheng, V. Reddy, P. Melnyk, and A. Mitra, "Three-dimensional fold of the human AQP1 water channel determined at 4 Å resolution by electron crystallography of two-dimensional crystals embedded in ice," *J. Mol. Biol.*, vol. 301, no. 2, pp. 369–87, 2000.

[7] R. Henderson, J. M. Baldwin, K. H. Downing, J. Lepault, and F. Zemlin, "Structure of purple membrane from *halobacterium halobium*: Recording, measurement and evaluation of electron micrographs at 3.5 Å resolution," *Ultramicroscopy*, vol. 19, pp. 147–178, 1986.

[8] M. Stewart, "Introduction to the computer image processing of electron micrographs of two-dimensionally ordered biological structures," *J. Electron. Micr. Tech.*, vol. 9, pp. 301–324, 1988.

[9] P. R. Smith and J. Kistler, "Surface reliefs computed from micrographs of heavy metal-shadowed specimens," *J. Ultrastruc. Res.*, vol. 61, pp. 124–133, 1977.

[10] ——, "Surface reliefs computed from micrographs of isolated heavy metal-shadowed particles," *J. Ultrastruc. Res.*, vol. 71, pp. 25–36, 1980.

[11] D. J. Müller, A. Engel, J. L. Carrascosa, and M. Vélez, "The bacteriophage $\phi 29$ head-tail connector imaged at high resolution with the atomic force microscope in buffer solution," *EMBO J.*, vol. 16, pp. 2547–2553, 1997.

[12] M. T. Chan, G. T. Herman, and E. Levitan, "Bayesian image reconstruction using image-modeling Gibbs priors," *Int. J. Imag. Syst. Tech.*, vol. 9, pp. 85–98, 1998.

[13] L. A. Amos, R. Henderson, and P. N. T. Unwin, "Three-dimensional structure determination by electron microscopy of two-dimensional crystals," *Prog. Biophys. Mol. Biol.*, vol. 39, pp. 183–231, 1982.

[14] R. Bracewell, *Two-Dimensional Imaging*. London, U.K.: Prentice-Hall, 1995.

[15] R. A. Crowther, R. Henderson, and J. M. Smith, "MRC image processing programs," *J. Struc. Biol.*, vol. 116, pp. 9–16, 1996.

[16] M. Yeager, V. M. Unger, and A. K. Mitra, "Three-dimensional structure of membrane proteins determined by two-dimensional crystallization. electron cryomicroscopy, and image analysis," *Methods Enzymol.*, vol. 294, pp. 135–180, 1999.

[17] E. R. Kunji, S. von Gronau, D. Oesterhelt, and R. Henderson, "The three-dimensional structure of halorhodopsin to 5 Å by electron crystallography: A new unbending procedure for two-dimensional crystals by using a global reference structure," in *Proc. National Academy of Science*, vol. 97, 2000, pp. 4637–4642.

[18] M. F. Schmid, R. Dargahi, and M. W. Tam, "SPECTRA: A system for processing electron images of crystals," *Ultramicroscopy*, vol. 48, pp. 251–264, 1993.

[19] D. A. Agard, "A least-squares method for determining structure factors in three-dimensional tilted-view reconstructions," *J. Mol. Biol.*, vol. 167, pp. 849–52, 1983.

[20] S. Matej, G. T. Herman, T. K. Narayan, S. S. Furuie, R. M. Lewitt, and P. E. Kinahan, "Evaluation of task-oriented performance of several fully 3D PET reconstruction algorithms," *Phys. Med. Biol.*, vol. 39, pp. 355–367, 1994.

[21] S. Matej and R. M. Lewitt, "Practical considerations for 3-D image reconstruction using spherically symmetric volume elements," *IEEE Trans. Med. Imaging*, vol. 15, pp. 68–78, Jan. 1996.

[22] R. Marabini, E. Rietsel, R. Schröder, G. T. Herman, and J. M. Carazo, "Three-dimensional reconstruction from reduced sets of very noisy images acquired following a single-axis tilt schema: Application of a new three-dimensional reconstruction algorithm and objective comparison with weighted backprojection," *J. Struct. Biol.*, vol. 120, pp. 363–371, 1997.

[23] R. Marabini, G. T. Herman, and J. M. Carazo, "3D reconstruction in electron microscopy using ART with smooth spherically symmetric volume elements (blobs)," *Ultramicroscopy*, vol. 72, pp. 53–65, 1998.

[24] C. O. S. Sorzano, R. Marabini, N. Boisset, E. Rietzel, R. Schröder, G. T. Herman, and J. M. Carazo, "The effect of overabundant projection directions on 3D reconstruction algorithms," *J. Struct. Biol.*, vol. 113, pp. 108–118, 2001.

[25] G. T. Herman, *Image Reconstruction From Projections: The Fundamentals of Computerized Tomography*. New York: Academic, 1980.

[26] ——, "Algebraic reconstruction techniques in medical imaging," in *Medical Imaging, Systems Techniques and Applications*, C. T. Leondes, Ed. Amsterdam, The Netherlands: Gordon and Breach , 1998, vol. 6, Computational Techniques, pp. 1–42.

[27] G. T. Herman, H. Levkowitz, S. McCormick, and H. K. Tuy, "Multilevel image reconstruction," in *Multiresolution Image Processing and Analysis*, A. Rosenfeld, Ed. New York: Springer-Verlag, 1984.

[28] R. M. Lewitt, "Alternatives to voxels for image representation in iterative reconstruction algorithms," *Phys. Med. Biol.*, vol. 37, pp. 705–716, 1992.

[29] S. Matej and R. M. Lewitt, "Efficient 3D grids for image reconstruction using spherically-symmetric volume elements," *IEEE Trans. Nucl. Sci.*, vol. 42, pp. 1361–1370, Oct. 1995.

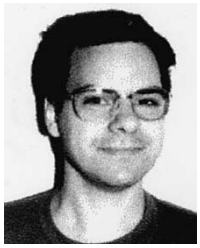
[30] R. M. Lewitt, "Multidimensional digital image representations using generalized Kaiser-Bessel window functions," *J. Opt. Soc. Amer. A*, vol. 7, pp. 1834–1846, 1990.

[31] S. S. Furuie, G. T. Herman, T. K. Narayan, P. E. Kinahan, J. S. Karp, R. M. Lewitt, and S. Matej, "A methodology for testing for statistically significant differences between fully 3D PET reconstruction algorithms," *Phys. Med. Biol.*, vol. 39, pp. 341–354, 1994.

[32] J. A. Reeds and L. A. Shepp, "Limited angle reconstruction in tomography via squashing," *IEEE Trans. Med. Imag.*, vol. MI-6, pp. 89–97, Feb. 1987.

[33] G. T. Herman, R. Marabini, J. M. Carazo, E. Garduño, R. Lewitt, and S. Matej, "Image processing approaches to biological three-dimensional electron microscopy," *Int. J. Imag. Syst. Tech.*, vol. 11, pp. 12–29, 2000.

[34] *International Tables for Crystallography (Space-Group Symmetry)*, vol. A, T. H. Hahn, Ed., Kluwer Academic, Dordrecht, The Netherlands, 1996.
[35] R. Marabini, I. M. Masegosa, S. M. M. C. San Martín, J. J. Fernández, C. V. L. G. de la Fraga, and J. M. Carazo, "Xmipp: An image processing package for electron microscopy," *J. Struct. Biol.*, vol. 116, pp. 237–240, 1996.
[36] J. M. Valpuesta, J. J. Fernández, J. M. Carazo, and J. L. Carrascosa, "The three-dimensional structure of a DNA translocating machine at 10 angstrom resolution," *Struct. Fold. & Des.*, vol. 7, pp. 289–296, 1999.
[37] A. A. Simpson, Y. Z. Tao, P. G. Leiman, M. O. Badasso, Y. N. He, P. J. Jardine, N. H. Olson, M. C. Morais, S. Grimes, D. L. Anderson, T. S. Baker, and M. G. Rossmann, "Structure of the bacteriophage $\phi 29$ DNA packaging motor," *Nature*, vol. 408, pp. 745–750, 2000.
[38] J. M. Carazo, "The fidelity of 3D reconstructions from incomplete data and the use of restoration methods," in *Electron Tomography. Three-Dimensional Imaging With the Transmission Electron Microscope*, J. Frank, Ed. New York: Plenum, 1992, pp. 117–166.
[39] S. B. Vardeman, *Statistics for Engineering Problem Solving*. Boston, MA: Academic, 1994.
[40] E. Dimmeler, R. Marabini, P. Tittmann, and H. Gross, "Correlation of topographic surface and volume data from three-dimensional electron microscopy," *J. Struc. Biol.*, vol. 136, pp. 20–29, 2001.



Roberto Marabini received the M.S. degree from the University Autonoma de Madrid, Madrid, Spain, in 1989 and the Ph.D. degree from the University of Santiago de Compostela, Santiago de Compostela, Spain, in 1995, respectively, both in physics.

He was previously with the BioComputing Unit of the National Center for BioTechnology, Spanish National Council of Scientific Research, Madrid, Spain, and then the University of Pennsylvania, Philadelphia, and the City University of New York from 1998 to 2002. At present, he is an Associate Professor in the "Escuela Superior Politecnica," University Autónoma de Madrid. His current research interests include inverse problems, image processing and high performance computing.

Carlos O. S. Sorzano was born in Málaga, Spain, in 1973. He received the B.Sc. degree in computer science, the M.Sc. degree (Hons.) in electrical engineering from ETSIT, Universidad Málaga, and the Ph.D. degree (Hons.) from the the BioComputing Unit, National Center for BioTechnology, Spanish National Council of Scientific Research, Madrid, Spain.

Currently, he is a Research Assistant in the Biomedical Imaging Group of the Swiss Federal Institute of Technology, Lausanne, Switzerland, and an Assistant Professor with the University of San Pablo–CEU, Madrid. His research interests include image processing, tomography, system identification, multiresolution approaches, and electron microscopy.

Samuel Matej received the M.Sc. degree (Hons.) in electrical engineering from the Slovak Technical University, Bratislava, Slovak Republic, in 1983, and the Ph.D. degree in tomographic image reconstruction from the Slovak Academy of Sciences, Bratislava, Slovak Republic, in 1988.

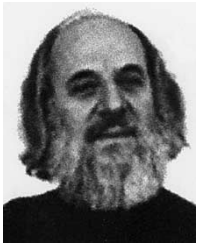
From 1988 to 1991, he was a Research Scientist at the Institute of Measurement Sciences, Slovak Academy of Sciences, Bratislava. In 1991, he joined the Medical Image Processing Group, Department of Radiology, University of Pennsylvania, Philadelphia, where he is currently a Research Associate Professor of Radiology. His research interests include methods of tomographic image reconstruction from projections in PET, SPECT, CT, MRI, and electron microscopy, with emphasis on data processing and reconstruction approaches for fully three-dimensional emission tomography.

José-Jesús Fernández received the M.S. and Ph.D. degrees in computer science from the University of Granada, Granada, Spain, in 1992 and 1997, respectively.

He was with at the BioComputing Unit of the Spanish National Center for BioTechnology, Spanish National Council of Scientific Research, Madrid, Spain. He became an Assistant Processor in October 1997 and, subsequently, Associate Professor of Computer Architecture at the University of Almería, Almería, Spain, in 2000. He is a member of the Supercomputing-Algorithms Research Group. His current research interests include high performance computing, image processing and tomographic reconstruction.

José M. Carazo received the M.Sc. degree (Hons.) from Granada University, Granada, Spain, in 1981, and the Ph.D. degree in molecular biology at the University Autónoma of Madrid (UAM), Madrid, Spain, in 1984.

He was at the IBM Madrid Scientific Center from 1981 to 1986, from where he moved to the New York State Health Department. In 1989, he set up the Biocomputing Unit of the then newly established National Center for Biotechnology, now the largest center of the Spanish National Council of Scientific Research (CSIC), located in the campus of the UAM. Currently, he is involved in the Spanish Ministry of Science and Technology as Deputy General Director for Research Planning while he remains engaged in his activities at the CNB, the UAM, the newly established Scientific Park of Madrid (PCM), and the start-up company Integromics. His current research interests focus in the development of new image processing approaches to resolve the architecture of molecular machines by hybrid techniques, including three-dimensional electron microscopy and X-rays diffraction.



Gabor T. Herman received the Ph.D. degree in mathematics from the University of London, London, U.K., in 1968, and honorary doctoral degrees from the University of Linkoping, Sweden, in 1989, the University of Szeged, Hungary, in 1998, and University of Haifa, Israel in 2000

He is a pioneer in the field of X-ray computerized tomography and the author of several books and over 100 articles, including many classic works in the field. He is recognized internationally for his major contributions to image processing and its biomedical applications. He was the leader of successful image processing groups at SUNY Buffalo and at the University of Pennsylvania and has garnered millions of dollars in research funding. He is currently a Distinguished Professor of Computer Science at The Graduate Center of the City University of New York.