VOLUMETRIC CONSTRAINTS IN 3D TOMOGRAPHY APPLIED TO ELECTRON MICROSCOPY

R. Marabini, G.T. Herman *

The Graduate Center
The City University of New York
New York
NY 10016
USA

C.O.S. Sorzano, J.M. Carazo †

Centro Nacional de Biotecnología-CSIC
Campus Universidad Autónoma de Madrid
28049 Madrid
Spain

ABSTRACT

3D Electron Microscopy aims at the reconstruction of density volumes corresponding to the mass distribution of macromolecules imaged with an electron microscope. There are many factors limiting the resolution achievable when this technique is applied to biomolecules: microscope transfer function, molecule flexibility, lack of projections from certain directions, unknown angular distribution, image noise, etc. In this communication we propose the use of *a priori* information such as particle symmetry, occupied volume, known surface, density nonnegativity and similarity to a known volume in order to improve the quality of the reconstruction. When a series expansion of the reconstructed volume is done, all these constraints are expressed as a set of equations which the expansion coefficients must satisfy. In this work, this equation set is specified and the effect of each one on the reconstruction of a realistic phantom is explored.

1. INTRODUCTION

3D Electron Microscopy is a powerful technique for structural studies of biological macromolecules due to the wide range of specimens that can be addressed, as well as the increasing resolutions achieved currently (50S and 70S subunits of E. Coli's ribosome at 7.5Å [1] and 11.5Å [2] respectively, Semliki Forest virus' capsid at 9Å[3], and GroEL at 11.5Å[4]). However, these high-resolution studies usually need a large number of projection images. In this work, we propose the incorporation of *a priori* information into the reconstruction process in a non-linear and adaptative way. This kind of *a priori* constraints are easily expressed in real space, and thus our reconstruction algorithm, ART+blobs [5], is specially suitable for dealing with them.

2. MATERIALS AND METHODS

ART with blobs is a series-expansion algorithm which assume that a volume $f(\mathbf{r})$ can be expressed as a linear combination of a set of basis funtions $b(\mathbf{r})$ properly placed in the space. For this work spherically symmetric functions termed blobs [6] are used as basis functions.

$$f(\mathbf{r}) \approx \sum_i x_i \cdot b(\mathbf{r} - \mathbf{r}_i)$$

Both, the blobs shape and their spatial distribution are fixed and the unknows are the coefficients x_i. These coefficients x_i are related to the experimental data (projections) by the equation:

$$\mathbf{p}_{\vec{\omega},\psi} = P_{\vec{\omega},\psi} \cdot \mathbf{x} \tag{1}$$

where $\vec{\omega}$ specifies the projection direction, ψ is the in-plane rotation, $\mathbf{p}_{\vec{\omega},\psi}$ is a vector formed by all the pixel values of the experimental projection in the direction $\vec{\omega}, \psi$, $P_{\vec{\omega},\psi}$ is the projection matrix for the reconstructed volume and $\mathbf{x}$ is the vector of all blob coefficients.

When available, other constraints different from the projection information can be imposed on the blob coefficients. Particularly, if we know the value of the function, v_j, at certain points, $\mathbf{r}_j \in K_V$, then we could state the following equation for every point at the set K_V

$$\sum_i x_i b(\mathbf{r}_j - \mathbf{r}_i) = v_j \tag{2}$$

or in matrix form

$$V \cdot \mathbf{x} = \mathbf{v} \tag{3}$$

where $\mathbf{v}$ is a vector with all known values and the element v_{ji} of the matrix V expresses the value of the blob with unit weight centered in $\mathbf{r}_i$ at the position $\mathbf{r}_j$, i.e., this is the matrix converting a blob volume into a voxel one. Expression

*Partially supported by NIH grant HL70472.

†Partially supported by grants BIO98-0761, BIO2001-1237 and TIC990361 from the "Comisión Interministerial de Ciencia y Tecnología" of Spain.

3 poses a new equation set, called *volumetric constraints*, that the reconstructed volume must satisfy. This equation set is also solved using ART after each single projection is presented to the reconstruction algorithm.

Symmetry

If a particle is symmetric then given any projection, automatically many other projections are also known as they should be exactly the same as the given one. Thus, a simple way of introducing symmetry is by presenting all the automatically generated projections to the reconstruction algorithm. It can be seen that the volumetric constraints are not needed to impose this constraint.

Known surface

There exist experimental microscopy techniques producing the image of the surface of a protein such as Metal Shadowing [7] or the Atomic Force Microscopy [8]. Furthermore, nonlinear image processing methods like thresholding, segmentation, and mathematical morphology can provide an approximation of the reconstruction volume. The information carried by such constraints force the reconstruction to be zero valued outside the known surface. Nothing is said about the density values inside the volume. This is modeled as a zero valued vector $\mathbf{v} = \mathbf{0}$ at the positions determined to be exterior to the surface K_S. With these definitions the volumetric constraints can be applied.

Density nonnegativity

In principle, the reconstructed volume corresponds to the mass density estimation of the protein under study. If a good normalization procedure is applied to the projection images, it can be guaranteed that the volume density at all points should be nonnegative. However, since a large amount of noise is present in the projection images it may turn out that certain positions within the volume have negative values. As this cannot be the case after the normalization, all those points with negative values are forced to be at least zero. This results in a set of known values $\mathbf{v} = \mathbf{0}$ for certain positions $K_D N$. The selection of known values is done on a nonlinear basis. Notice that the set $K_D N$ changes from iteration to iteration since the points with negative values usually change. In this sense this constraint is adaptive and nonlinear.

Occupied volume

Experimentalists can determine the molecular weight of the complex under study via biochemical experiments or some estimation based on the protein aminoacid sequence. This weight can easily be translated into an estimated number of voxels that the reconstructed volume should occupy. In fact, this is the most used criterion for segmenting reconstructions in the field: this number of voxels determines a density threshold used for segmentation. If projections have been properly normalized, a similar reasoning to that for the density nonnegativity can be used and all values below that threshold $K_O V$ could be set to zero. This is again a nonlinear and adaptive constraint.

Similarity to a known volume

It is known that the reconstruction algorithm based on ART tends to minimize the least squares distance between the iterative solution at each stage and the initial one [9]. Usually a zero valued initial volume is provided as starting point to the iterative algorithm expressing the uncertainty about the correct reconstructed volume and trying to minimize the variance of the resulting reconstruction, with the intention of minimizing the noise. However, there are many situations where a good enough reconstruction is available, which can be used as starting point for the iterations. The volumetric constraints can be used alone to convert volumes expressed in voxels to blobs. Now, $\mathbf{v}$ stands for all values in the initial solution and K_V is the support where this solution is defined.

3. RESULTS

To test the efficacy of the proposed constraints 500 projection images were simulated from an atomic model of bacteriorhodopsin [10] (see Figure 1 left). The collection geometry simulates that obtained in random conical tilt [11, Chap. 5]. A signal to noise ratio of 1/3 was simulated. Figure 1 shows an isosurface rendering accounting for 100% of the mass of the outputs for the reconstruction without and with all the constraints. Figure 2 represent some central slices of the phantom and the two reconstructions. The reconstructed volumes have not been filtered, symmetrized nor masked to better show the amount of noise present, and for this reason the resolution achieved in each case is so low.

Figure 3 represents the achieved resolution (measured with Fourier Shell Correlation [11, Chap. 5]) when every constraint is applied separately, when none is applied and when all are applied at the same time. In this last case, the resolution increases from 83Å or 52Å to 44Å.

4. DISCUSSION

The experiment carried out shows that the constraints imposed nearly double the resolution achieved when no postprocessing is performed and that these constraints can supply enough information to recover information at frequencies that were greatly attenuated. There is a dip in the Fourier Shell Correlation of the raw reconstruction

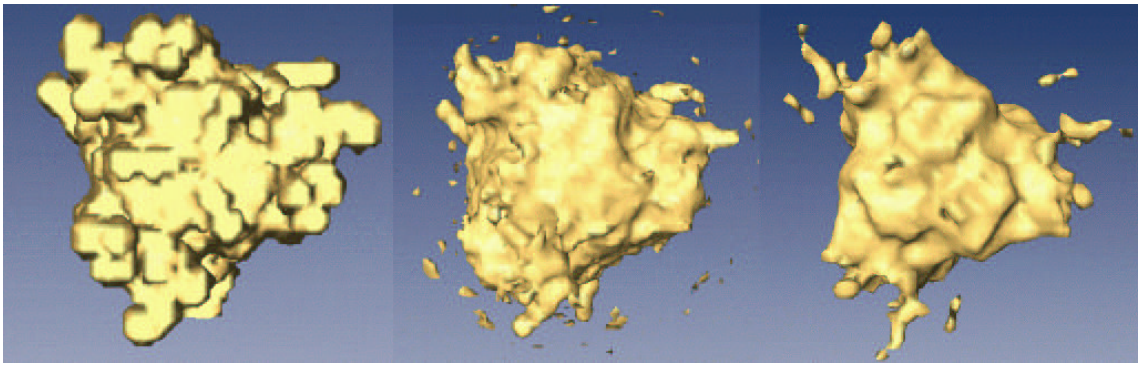

Fig. 1. Isosurface containing 100% of the bacteriorhodopsin mass for the atomic phantom (left), the reconstruction performed with ART+blobs without constraints (middle), and the reconstruction performed with ART+blobs with constraints (right).

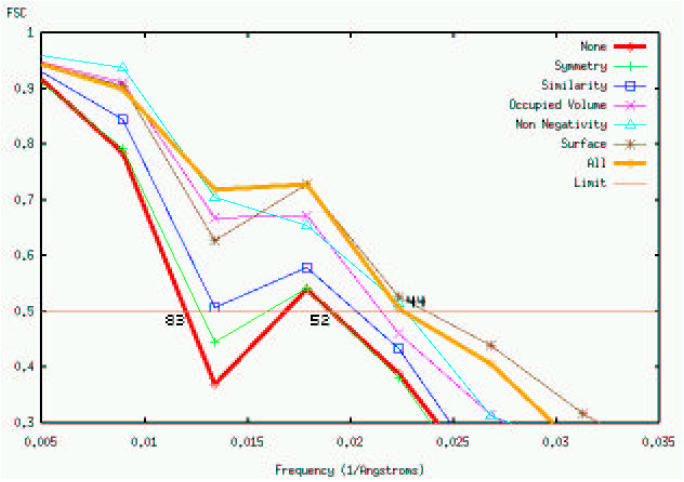

Fig. 3. Fourier Shell Correlation (FSC) for the reconstructions compared with the bacteriorhodopsin phantom when the different constraints are applied. Constraints are ordered in the legend according to their ability to increase the resolution. The FSC for the cases when all and none of the constraints are applied are in bold.

(the one done without constraints) that is mantained in the reconstructions with symmetry, the macromolecule surface or when an initial solution is provided and does not appear in reconstructions with nonnegativity or the occupied volume. This implies that the nonlinearity of the latter two constraints is able to supply the right information at the affected frequencies.

The different constraints supply different amounts of information. It seems that the least informative is symmetry. A possible explanation for this is that the symmetry element (a rotational axis) is aligned with the missing cone generated by the random conical tilt geometry, and for this reason the automatically generated projections are not adding information to what had already been measured.

The next least informative is the constraint providing an initial point to the reconstruction algorithm. In this work, the reconstruction produced without constraints has been used as initial guess. Thus, the effect shown is as the one of iterating twice on the projection set. Nevertheless, disregarding the origin of the initial solution, it is clear that a good approximation to the problem solution helps the iterative algorithm to find a better reconstruction.

The next more informative constraints are the occupied volume and the nonnegativity of the reconstructed densities. These two nonlinear constraints are similar in performance, although nonnegativity seems to be a little bit stronger. Their nonlinear nature seems to be able to enhance information that was greatly attenuated as is the case of the dip observed around 70Å in the Fourier Shell Correlation (see Figure 3).

The macromolecular surface is a powerful constraint as it gives much information regarding the complex boundary. In this work the exact molecular surface has been provided since the protein is absolutely known. There are experimental techniques [7, 8] which provide the surface of a macromolecular complex with a high accuracy. The resolution

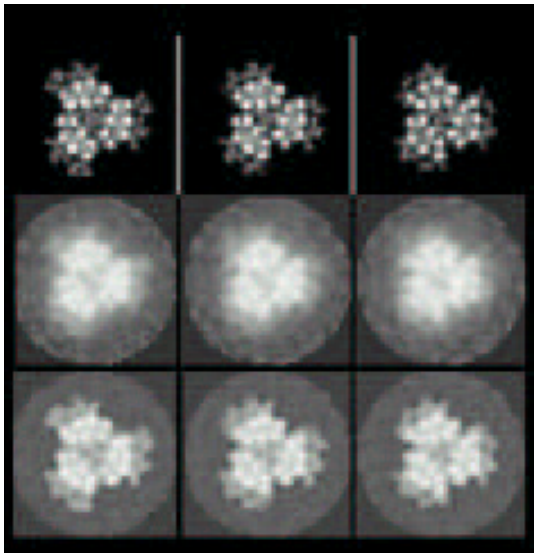

Fig. 2. Sample slices of the bacteriorhodopsin phantom (top row), the reconstruction without constraints (middle row) and the reconstruction with constraints (bottom row).

achieved by the reconstruction with only the surface constraint is the highest one although in general it is not the best reconstruction, as is shown by the rest of the FSC curve.

When all constraints are applied at the same time, a kind of consensus among them all is produced. Although the FSC in this case is not the highest one at all frequencies, it can be regarded as the best one as a whole.

5. CONCLUSION

We have extended the ART+blobs reconstruction algorithm to include volumetric constraints such as symemtry, occupied volume, known surface, density nonnegativity and similarity to a known volume. While they all have shown to be informative, the nonlinear constraints appear to be specially powerful. The combination of all the constraints greatly improves the resolution achieved.

6. REFERENCES

[1] F. Mueller, I. Sommer, P. Baranov, R. Matadeen, M. Stoldt, J. Wohnert, M. Gorlach, M. van Heel, and R. Brimacombe, "The 3D arrangement of the 23 S and 5 S rRNA in the Escherichia coli 50 S ribosomal subunit based on a cryo-electron microscopic reconstruction at 7.5 Å resolution," *J. Molecular Biology*, vol. 298, no. 1, pp. 35–59, Apr. 2000.

[2] I. S. Gabashvili, R. K. Agrawal, C. M. Spahn, R. A. Grassucci, D. I. Svergun, J. Frank, and P. Penczek, "Solution structure of the *E. coli* 70S ribosome at 11.5 Å resolution," *Cell*, vol. 100, pp. 537–49, 2000.

[3] E. J. Mancini, M. Clarke, B. E. Gowen, T. Rutten, and S. D. Fuller, "Cryo-electron microscopy reveals the functional organization of an enveloped virus, Semliki Forest virus," *Molecular Cell*, vol. 5, no. 2, pp. 255–266, 2000.

[4] Steven J. Ludtke, Joanita Jakana, Jiu-Li Song, David T. Chuang, and Wah Chiu, "A 11.5 Å single particle reconstruction of GroEL using EMAN," *J. Molecular Biology*, vol. 314, pp. 253–262, Nov. 2001.

[5] 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.

[6] R. M. Lewitt, "Multidimensional digital image representations using generalized Kaiser-Bessel window functions," *J. Optical Society of America A*, vol. 7, pp. 1834–1846, 1990.

[7] K. G. Fuchs, P. Tittmann, K. Krusche, and H. Gross, "Reconstruction and representation of surface data from two-dimensional crystalline biological macromolecules," *Bioimaging*, vol. 3, pp. 12–24, 1995.

[8] A. Engel and D. J. Muller, "Observing single biomolecules at work with the atomic force microscope.," *Nature Structural Biology*, vol. 7, no. 9, pp. 715–718, Sept. 2000.

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

[10] T. A. Ceska, R. Henderson, J. M. Baldwin, F. Zemlin, E. Beckmann, and K. Downing, "An atomic model for the structure of bacteriorhodopsin, a seven-helix membrane protein," *Acta Physiol. Scand. Suppl.*, vol. 607, pp. 31–40, 1992.

[11] J. Frank, *Three Dimensional Electron Microscopy of Macromolecular Assemblies*, Academic Press, San Diego, CA, USA, 1996.