A Method for Refining the Geometry of Images Collected by the Random Conical Tilt Scheme Based on Common Lines

Luis Gerardo de la Fraga § Gabor T. Herman ‡

José María Carazo §

§ BioComputing Unit, Centro Nacional de Biotecnología-CSIC,
Campus Universidad Autónoma, 28049 Madrid, Spain.

‡ Medical Image Processing Group, Department of Radiology,
University of Pennsylvania, Philadelphia, PA 19104, USA.
E-mail: fraga@cnb.uam.es

Abstract

A group of N two-dimensional projection images collected by the random conical tilt scheme [1] is characterized by φ as the tilt angle and θ_n as the rotational angles. In this communication we present a new method to refine the estimators of these two angles. We use an approach based on common lines. The method is demonstrated on one synthetic model of DnaB protein.

1 Introduction

To perform a three-dimensional reconstruction from a group of two-dimensional projections we need information regarding the localization of each projection relative to a coordinate system fixed to the volume. Given a 2-D projection, its localization is expressed by that we know as the Euler angles (we refer to these as the *tilt* and the *rotation*) and by the two coordinates of the location in the projection plane onto which the origin of the volume coordinate system is projected.

The random-conical reconstruction scheme [1, 2] is much used for the purpose of estimate the localization, the Euler angles and origin coordinates, for projection images taken with the electron microscope, in order to perform a 3D reconstruction of biological macromolecules. This technique uses a pair of micrographs (the first tilted and the second untilted) of a specimen containing many particles, prepared in such a way that all the particles assume a preferred orientation with respect to the specimen support. If this is performed successfully, each particle image seen in the tilted micrograph is a projection with a fixed tilt, but unknown rotation. The untilted view (in which the rotation of the volume around the common axis of the particles is reflected by rotation by the same angle in the projection) can be used to estimate the rotation. Also, by averaging the unrotated versions of the particles in the untilted view, we can get a high quality estimate of the untilted projection and consequently a good estimate of the projection of the origin of the volume coordinate system. (Here we may assume that this origin is at the centroid of the volume and make use of the fact that, if noise is negligible, then the centroid of the projection is a good estimate of the centroid of the volume.) Finally, one may estimate the coordinates of the projection of the origin in the tilted projection images by the the so-called *cosine stretching* [1, 3] procedure: the tilted projection is stretched by the factor $1/\cos(\varphi)$ in the direction perpendicular to the tilt axis, where φ is the tilt angle, and cross-correlated with the averaged untilted projection. The random-conical reconstruction scheme has found wide popularity because, in principle, it solves the problem of finding the relative orientations of the different projections.

However, the methodology described in the previous paragraph produces a less than perfect estimation of the geometrical parameters of the projections with respect to the assumed fixed coordinate system attached to the volume for three reasons. (i) The tilt angle in reality is not identical for all particles, but encompasses a range, of possibly as much as 20°, due to the support not being completely flat. (ii) The rotation angles are determined from the untilted image, in which the particles have been previously exposed (while taking the tilted image) to the electron beam and may have been damaged. (iii) The location of the projection of the origin of the assumed coordinate system (attached to the particles) may be incorrectly determined by the cosine stretching procedure for volumes which are not very flat.

In this work we present a methodology to solve problems (i) and (ii) listed above. To correct (iii) we developed a new method [4] that solve the problem (as well as we can), by identifying in the projection image the location onto which the volume centroid projects purely from the projection image. Although angular deviations could be as far as 20 degrees on the tilt angle, here we are interested in obtaining the "best" particles that are according to one random conical tilt geometry, and generate, with all of them, the "best" initial 3D reconstruction. All the other images rejected by the method can be incorporated to the initial model using a "refinement method" such as [5, 6, 7] using the initial 3D reconstruction as a reference model.

2 Method description

We suppose that we have a group of I two-dimensional projection images whose orientation is described by ϕ as the tilt angle and θ_i as the rotational angles. Our problem is to identify ϕ and find a consistent subset of θ_i.

It can be proven that between any pair of two-dimensional projections we have a common line, a fact exploited for asymmetrical particles in [8]. Furthermore, for the random conical tilt geometry a further derivation proves that the common-line is at an angle ρ, over the first image, and $-\rho$ over the second one, so that

$$\tan\rho = \frac{\cos\phi(1-\cos\theta)}{\sin\theta} \tag{1}$$

where $\theta_1 = 0$, and $\theta_2 = \theta$, for each pair of images.

We use the quadratic distance between unidimensional projections

$$\mathrm{D}(\rho, j, k) = \sum_{r \text{ rays}} [\mathrm{p}_j(\rho - \pi/2, r) - \mathrm{p}_k(-\rho - \pi/2, r)]^2 \tag{2}$$

of images j and k, as the function to locate the common-line by searching around one estimated point calculated using the approximated initial values for all the angles (the initial value of ϕ is taken from the reading of goniometer's microscope, and the initial values of θ's are taken by the alignment of the "zero-tilt-views").

We assume the next: a) Range of the tilt angle: $0 < \phi < \pi/2$. b) Range of rotational angles: $0 < \theta_i < 2\pi$, and c) Range of ρ angle: $-\pi/2 < \rho \leq \pi/2$. Our method consist of the following steps. We discretize all angles to delta units (du), 1 du $= \pi/2048$:

STEP 1:

For each pair of images (j, k), calculate $\rho j, k$, from Eq. (1), using the initial angles values of ϕ and θ_i:

$$\rho_{j,k} = \rho(\phi, \theta_j, \theta_k), \quad 0 \leq j < I-1, \quad j+1 \leq k < I$$
$$\rho_{k,j} = -\rho_{j,k}$$

STEP 2:

We will check if the common-line for all possible pairs of images can be obtained. Create a table (an upper triangular matrix) of size $I(I-1)/2$. For each pair of images we will calculate five points around $\pm 64du$ (approx. $\pm 5.6°$):

$$D(\rho_{jk} - 64\delta, j, k), D(\rho_{jk} - 32\delta, j, k), D(\rho_{jk}, j, k),$$
$$D(\rho_{jk} + 32\delta, j, k), D(\rho_{jk} + 64\delta, j, k)$$

if the smaller value, from these five, is at one of the extreme locations, we enter a "black mark" on the created table at location (j, k).

STEP 3:

Remove the image with more black marks by removing the corresponding row or column in the table. If there are equal number of black marks in a row and column, remove the row. Update I value.

STEP 4:

Now we can find the localization where the common-line is. For each pair of images (j, k), search for the minimum of function D, Eq. (2) (the range of searching is $\pm 64du$ around the initial estimated point). The $\mathring{\rho}_{jk}$ corresponds to the minimum value. Calculate

$$M = \frac{2}{I(I-1)} \sum_{j=0}^{I-2} \sum_{k=j+1}^{I-1} \left(\mathring{\rho}_{jk} - \rho_{jk}\right)^2 \tag{3}$$

STEP 5:

For each pair of images (j, k) estimate the values $\phi_{j,k}$ using Eq. (1), at the corresponding $\mathring{\rho}_{jk}$:

$$\phi_{jk} = \phi(\mathring{\rho}_{jk}, \theta_j, \theta_k)$$

STEP 6:

Find the three most typical images, by eliminating each time the most atypical image as below: Remove image l that maximize:

$$\left| \frac{2}{N(N-1)} \sum_{j \neq k} \phi_{jk} - \frac{2}{(N-1)(N-2)} \sum_{\substack{j \neq k \\ j \neq l \\ k \neq l}} \phi_{jk} \right| \tag{4}$$

for, at beginning, $N = I$, $0 \leq j < N-1$, $j < k < N$. We repeat this process until there are three images left. In the same process images were ordered by "typicality" and now the most atypical image correspond to $i = I$. Update ϕ value to:

$$\mathring{\phi} = \varphi_{12} \quad \text{or} \quad \varphi_{13} \quad \text{or} \quad \varphi_{23} \tag{5}$$

In this point, the three most typical images and the associated $\mathring{\rho}(j, k)$ (there are three of them) uniquely determine ϕ, $\theta_2 - \theta_1$, $\theta_2 - \theta_1$. Finally, by insisting that θ_1, θ_2, and θ_3 matches to the initially estimated values, we now have a ϕ, θ_1, θ_2, and θ_3, which are updates of the values initially estimated. We take any estimated value of ϕ because the three most typical images must have their θ values correct.

STEP 7:

Suppose now that we have a set of estimated of ϕ, θ_1, $\theta_2, \ldots, \theta_n$. For a new image $k = n + 1$ do

1. Calculate [using Eq. (1)]
$$\overline{\theta}_{n+1} = \frac{1}{n} \sum_{j=1}^{n} \theta_{n+1}(\mathring{\phi}, \mathring{\rho}_{jk}, \theta_j) \tag{6}$$
2. Calculate new estimated values of $\phi_{j,k}$ (as in STEP 5).
3. Update $\mathring{\phi}$ value:
$$\mathring{\phi}_{n+1} = \frac{m_n \mathring{\phi}_n + \sum_{j=1}^{n} \phi_{jk}}{m_n + n}, \qquad \text{para } n > 3 \tag{7}$$
where
$$m_n = \frac{(n-1)(n-2)}{2}$$
The initial $\mathring{\phi}_3$ was is calculated using Eq. (5).
4. Recalculate new values $\rho_{j,k}$ (same that STEP 1) but using the new refined values $\mathring{\phi}$, Eq. (7), and θ_k, Eq. (6).
5. Calculate M' in the same manner that Eq. 3 (STEP 4), but now over j. If $M' > M$ then
STOP the process
Otherwise continue with the next image ($n \leftarrow n + 1$).

3 Stability considerations

We can rewrite Eq. (1) as:

$$\cos\phi = \tan\rho \frac{\sin\theta}{1-\cos\theta}$$
$$\cos\phi = \tan\rho \, g(\theta) \tag{8}$$

In Eq. (8) we can see that the product of functions $\tan\rho$ and $g(\theta)$ must have the same restriction that the cosine function, i.e., $|\tan\rho \, g(\theta)| \leq 1$.

In our method, on STEP 5, we estimate new values $\phi_{j,k}$, for each pair or images. These $\phi_{j,k}$ values depend on the location of the common-line, given by the value of $\mathring{\rho}_{j,k}$ angle, and the noisy values of θ_j and θ_k. We could get undefined values of $\phi_{j,k}$, in other words, the product $\tan\rho \, g(\theta)$ could gives us a value greater than 1, for the cases in which the difference $\theta_k - \theta_j$ is close to zero or close to 2π. In fact, this indetermination produced, in out first runs of our method, errors in the built-in acos function provided by the compiler.

The explanation of this behavior is that the function $g(\theta)$ asymptotically tends to infinity, therefore, it is unstable, for θ values close to 0 or 2π radians. We can see the graphic of the $g(\theta)$ function on Fig. 1.

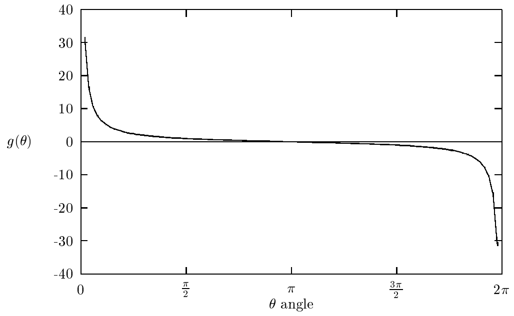


Figure 1: Graphic of function $g(\theta)$

3.1 Correction to the process

Due the unstabilities on the calculus of $\phi_{j,k}$, the STEP 5 in our method is modified to:

STEP 5:

For each pair of images (j,k) estimate the values $\phi_{j,k}$ using Eq. (1), at the corresponding $\mathring{\rho}_{jk}$:

$$\phi_{jk} = \phi(\mathring{\rho}_{jk}, \theta_j, \theta_k)$$

If for any pair, ϕ can not be calculated, do:

$\phi_{j,k} =$ UNDEFINED

If for any new value of $\phi_{j,k}$ occurs

$$\left| \mathring{\phi} - \phi_{jk} \right| > \Phi$$

do:

$\phi_{j,k} =$ UNDEFINED

The same considerations are taken on STEP 7.2.

STEP 6:

We sort the images according to a "distance map" to the average of $\phi_{j,k}$, $\overline{\phi}$. We calculate $\overline{\phi}$ over all defined values of $\phi_{j,k}$, and we create a table with distance values defined as:

$$d_{j,k} = \phi_{j,k} - \overline{\phi} \tag{9}$$

For any undefined $\phi_{j,k}$, we add a big value, i.e. 10^4. We define for each image i:

$$c_i = \sum_{j=i}^{N-2} |d_{i,j}| + \sum_{j=0}^{i-1} |d_{j,i}| \tag{10}$$

Now, the three most typical images are equivalent to the three first images with the smallest c_i.

STEP 7:

We introduce a slight modification in this STEP. The STOP condition is changed to the following:

If the condition $M' > M$ is valid, we do not stop all the process. Instead, the image that produce the stop condition is sent to the end, and it is taken now as the most atypical image.

We can see a flowchart of all the process in Fig. 2.

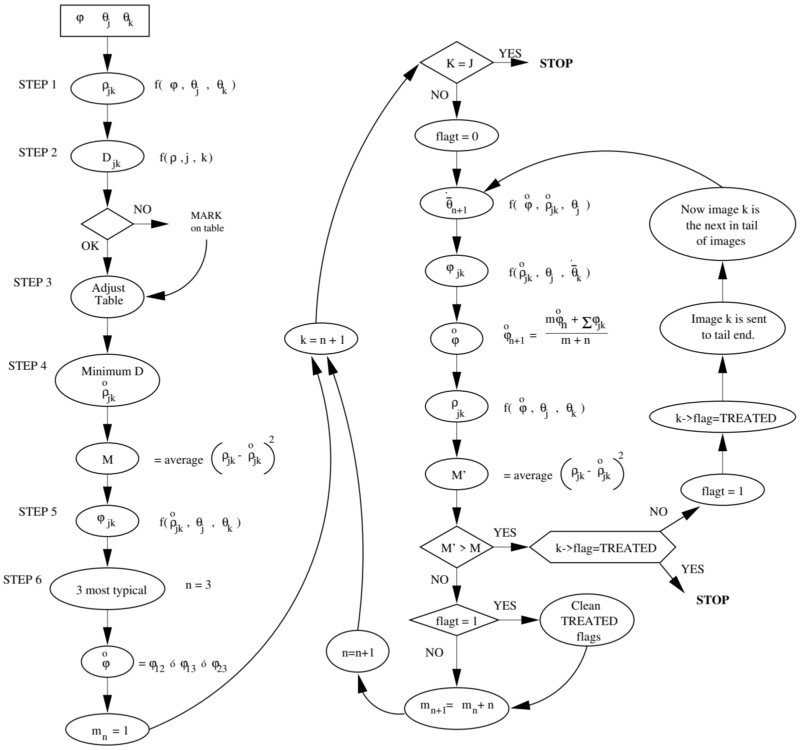


Figure 2: Flowchart of all the process

4 Experiment

We test our method with a phantom, a synthetic model of DnaB protein [9]. We can see several surface views of our phantom in Fig. 3.

We take 65 projections of the phantom at $\phi = 50°$, in tilt angle, and rotational angles $n \cdot 2\pi/(3 \cdot 65)$, $\quad 0 \leq n < 65$ We added a zero-mean Gaussian noise to each projected image to get fractional noise levels [10] of 0.05, 0.10, and 0.20 The level of noise added to each projection is expressed as the ratio between the standard deviation of the noise added and the average intensity range for all simulated images. We can see several examples of these projections in Fig. 4.

The common-line is at $\phi = 16°$ (over the first image) for a pair of projections taken at $\phi = 50°$, $\theta_1 = 0°$ for the first image, and $\phi = 50°$, $\theta_2 = 48°$ for the second one. This composition is shown in Fig. 5. The positive ρ angle is on the first image. The 1D projections of the two images are taken at angles $\rho + \phi$, over the first, and at $-\rho + \phi$ over the second one, and for the calculus of D function, Eq. (2), the quadratic difference between these two 1D projections is calculated.

The shape of the function D, Eq. (2), calculated with the two first images in Fig. 4, at different fnl, are shown in Fig. 6 at fnl=0.00, Fig. 7 at fnl=0.05, Fig. 8 at fnl=0.10, and in Fig. 9 at fnl=0.20.

We can see the results of one pass of our method over the 65 projections in Figs. 10.

5 Conclusion

In short, the method described here is a specially derived application of the principles of common lines to the specifics of a random conical tilt data acquisition geometry. It is based on pair-wise comparisons, and it has been shown to be reliable for images with a fractional noise level (fnl) [10] up to 0.2. Cryo-electron microscopy images are still noisier than the ones this method can yet handle, and for that application others methods, such [5] and [6, 7, 11], should be used. However, for a wide range of other preservation and imaging modes, the method in this contribution could indeed be very valuable.

Acknowledgments

This work was supported in part by grant BIO95-0768 from the Spanish CICYT, while the collaboration between Spain and USA was made possible by NATO grant CRG 960070.

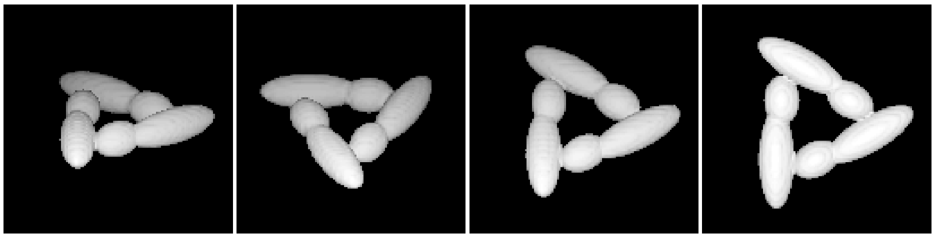


Figure 3: Surface views of our DnaB phantom

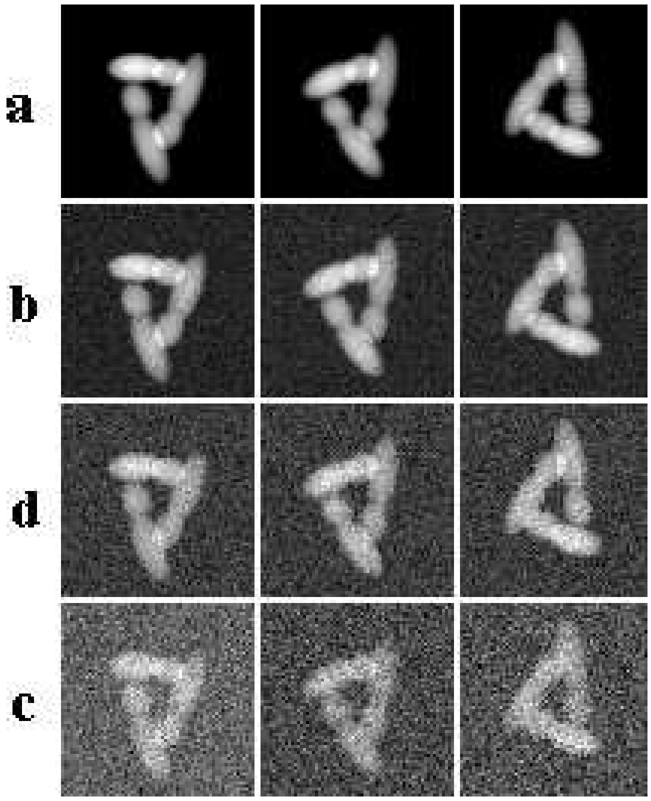


Figure 4: Examples of projected images, a) without noise. b) The same images than a) but with a noise at $fnl = 0.05$. c) As before but at fnl= 0.10. d) The same images but increasing the fnl up to 0.20

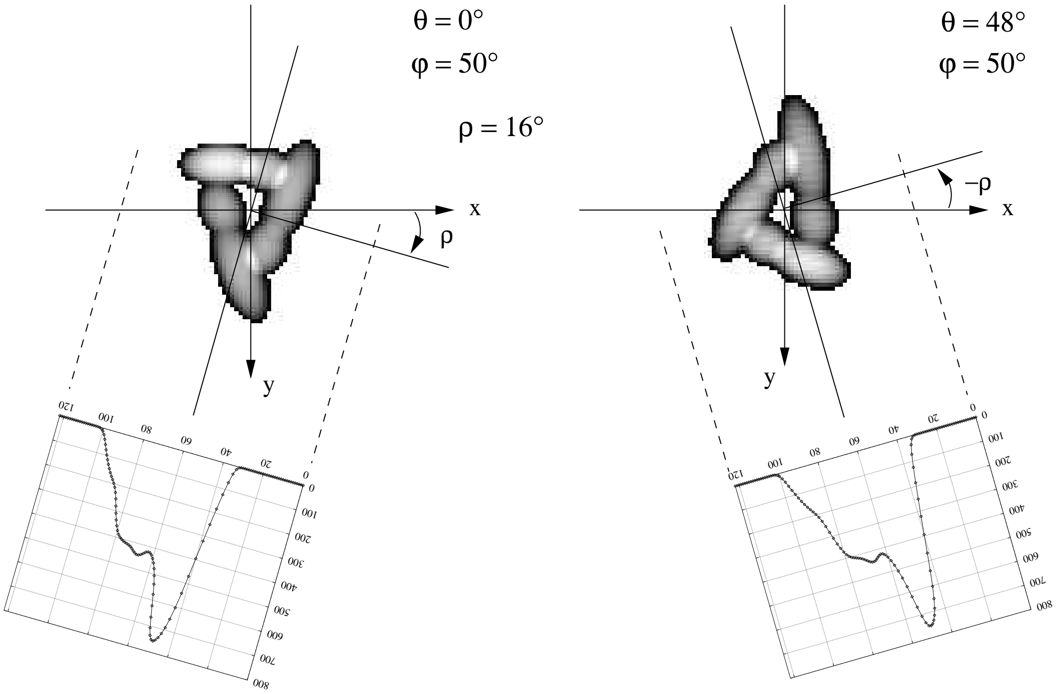


Figure 5: Visualization of the common tilt axis (the common-line in the Fourier space) for two images taken according to a conical tilt geometry

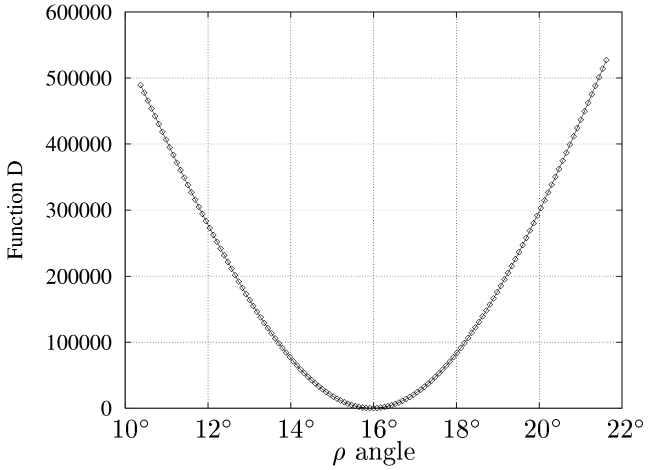


Figure 6: The function D, in Eq. 2, calculated for the fist two images in Fig. 4.a, corresponding to fnl=0.00

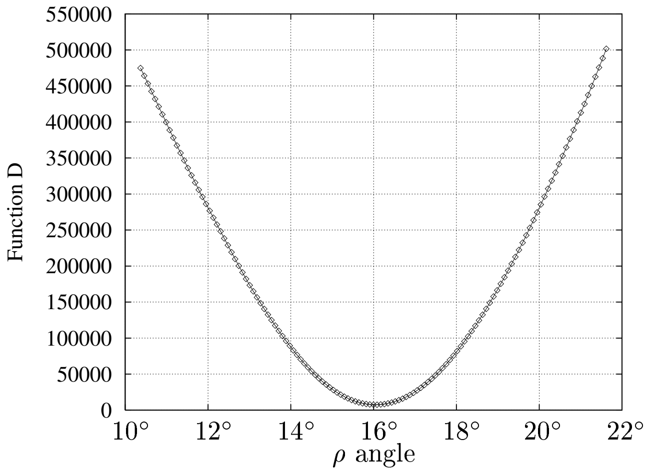


Figure 7: The function D, in Eq. 2, calculated for the first two images in Fig. 4.b. It corresponds to the images at fnl=0.05

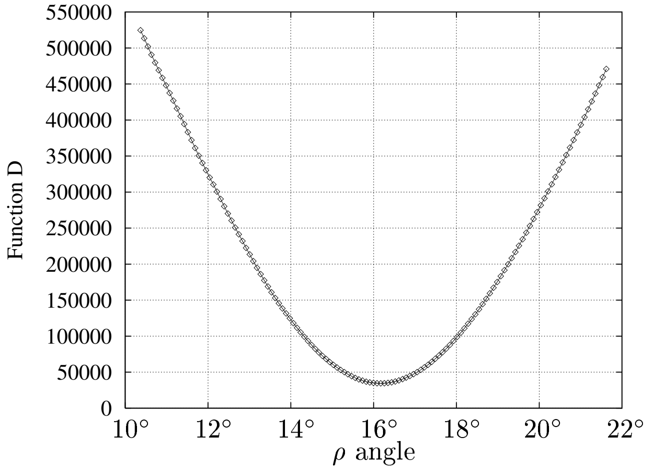


Figure 8: The function D, in Eq. 2, calculated for the first two images in Fig. 4.c. It corresponds to the images at fnl=0.10

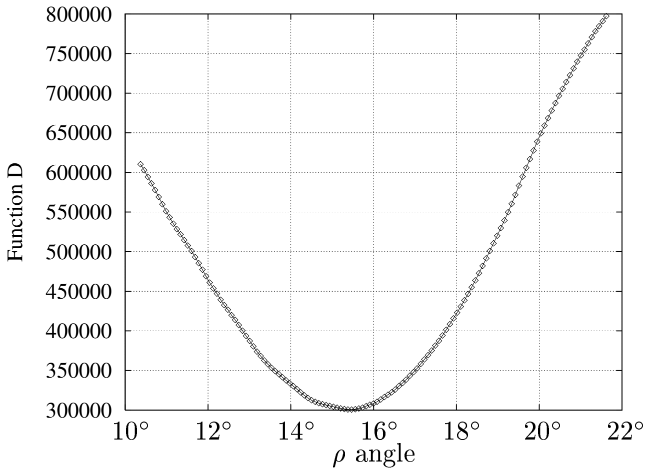
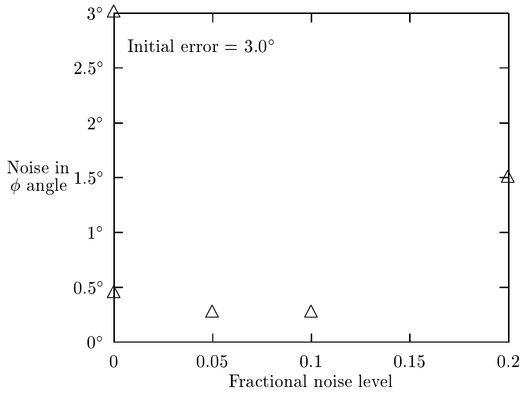
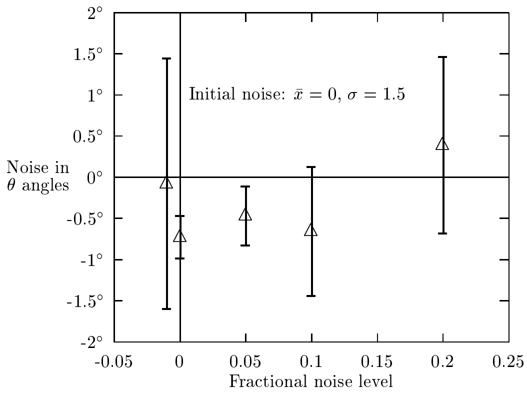


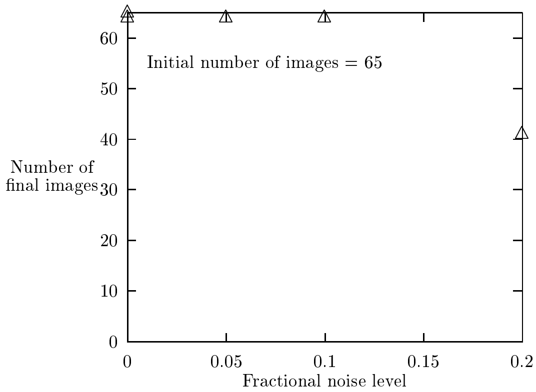
Figure 9: The function D, in Eq. 2, calculated for the two first images in Fig. 4.d. It corresponds to the images at fnl=0.20



(a)



(b)



(c)

Figure 10: Results of one pass of our method over the 65 projections at several values of fnl. a) Error in the tilt angle (ϕ). b) Mean and standard deviation of the error in the rotational angles. c) Final number of images.

References

[1] M. Radermacher. The three-dimensional reconstruction of single particles from random and non-random tilt series. *J. Electron Microsc. Tech.*, 9:359–394, 1988.

[2] O. Lambert, N. Boisset, J.-C. Taveau, G. Préaux, and J. N. Lamy. Three-dimensional recontruction of the α_D and β_C-hemocyanins of *helix pomatia* from frozed-hidrated specimens. *J. Mol. Biol.*, 248:431–448, 1995.

[3] J. M. Carazo and J. Frank. Three-dimensional matching of macromolecular structures obtained from electron microscopy. *Ultramicroscopy*, 25:13–22, 1988.

[4] L.G. de la Fraga, J.M. Carazo, and G.T. Herman. A method for obtaining a low resolution 3d reconstruction from very noisy projections. In *Proc. VII National Symposium on Pattern Recognition and Image Analysis*, volume 1, pages 347–352, 1997.

[5] P. A. Penczek, R. A. Grassuni, and J. Frank. The ribosome at improved resolution: new techniques for for merging and orientation refinement in 3D cryo-electron microscopy of biological particles. *Ultramicroscopy*, 53:251–270, 1994.

[6] M. Radermacher. Three-dimensional reconstruction from random projections: Orientational alignment via Radon transforms. *Ultramicroscopy*, 53:121–136, 1994.

[7] T. S. Baker and R. H. Cheng. A model-based approach for determining orientations of biological macromolecules imaged by cryoelectron microscopy. *J. Struct. Biol.*, 116:120–130, 1996.

[8] M. van Heel. Angular reconstitution: a posteriori assignment of projection direction for 3D reconstruction. *Ultramicroscopy*, 21:111–124, 1987.

[9] M. C. San Martín, M. Radermacher, B. Fedtke, A. Engel, C. S. Miles, N. E. Dixon, and J. M. Carazo. Three-dimensional reconstructions from cryo-electron microscopy images reveal an intimate complex between helicase DnaB and its loading partner DnaC. *Structure*, 6:501–509, 1998.

[10] N. A. Farrow and F. P. Ottensmeyer. A posteriori determination of relative projection directions of arbitrarily oriented macromolecules. *J. Opt. Soc. Am. A*, 9(10):1749–1760, 1992.

[11] M. Radermacher. *Scanning Microsc. Int. Suppl.*, 1998.