A novel approximation method of CTF amplitude correction for 3D single particle reconstruction

Jiang, L.H.; Liu, Z.F.; Georgieva, D.; Kuil, M.E.; Abrahams, J.P.

Ultramicroscopy 110 (2010) 350–358

Contents lists available at ScienceDirect

Ultramicroscopy

journal homepage: www.elsevier.com/locate/ultramic

A novel approximation method of CTF amplitude correction for 3D single particle reconstruction

Linhua Jiang*, Zunfeng Liu, Dilyana Georgieva, Maxim E. Kuil, Jan Pieter Abrahams

Department of Biophysical Structural Chemistry, Leiden Institute of Chemistry, Leiden University, Einsteinweg 55, 2333 CC Leiden, The Netherlands

ARTICLE INFO

Article history:
Received 17 March 2009
Received in revised form
23 November 2009
Accepted 19 January 2010

Keywords:
Single particle
3D reconstruction
CTF correction
Image processing software

ABSTRACT

The typical resolution of three-dimensional reconstruction by cryo-EM single particle analysis is now being pushed up to and beyond the nanometer scale. Correction of the contrast transfer function (CTF) of electron microscopic images is essential for achieving such a high resolution. Various correction methods exist and are employed in popular reconstruction software packages. Here, we present a novel approximation method that corrects the amplitude modulation introduced by the contrast transfer function by convoluting the images with a piecewise continuous function. Our new approach can easily be implemented and incorporated into other packages. The implemented method yielded higher resolution reconstructions with data sets from both highly symmetric and asymmetric structures. It is an efficient alternative correction method that allows quick convergence of the 3D reconstruction and has a high tolerance for noisy images, thus easing a bottleneck in practical reconstruction of macromolecules.

1. Introduction

The last decade saw a substantial increase in the number of 3D structures determined by single particle cryo-EM reconstruction and the resolution of these reconstructions (~4–10 Å) is starting to approach a level which allows atomic interpretation of the structures (see reviews by Zhou [1] and Chiu et al. [2]). The development of procedures for accurate CTF estimation and correction of the measured image data was essential. The instrumental aberration problem that affects electron microscopy images was recognized early [3,4] and must be corrected for to allow the resolution to be extended beyond the first zero of the oscillating contrast transfer function (CTF). Multiple reconstruction software packages were adapted in this fashion to allow constructing high resolution 3D models e.g. IMAGIC [5,6], SPIDER [7,8], XMIPP [9,10], EMAN [11], IMIRS [12] and others. About seven parameters (depending on the CTF model used) need to be determined in the CTF estimation for an accurate approximation. These parameters are subsequently used in the CTF correction procedure. The quality of the final 3DEM model relies on accurate CTF estimation and correction. This makes CTF estimation and correction one of the most delicate problems in 3D single particle reconstruction.

For CTF estimation, a number of semi-automatic tools are available (e.g. [13–16]). There are also fully automatic CTF estimation tools, based on different methods, e.g. ARMA models of Xmipp [17]; ACE: Automated CTF Estimation [18]; automatic CTF estimation based on multivariate statistical analysis [19]. Here we describe a new method for correcting images optimally when (initial) estimates of the CTF parameters are available.

According to the theory [4,20,21], the image measured in TEM normally can be described in Fourier space as a function of the spatial frequency vector s by

$$M(s) = CTF(s)F(s) + N(s) \quad (1)$$

$M(s)$ is the Fourier transform of the measured image. $CTF(s)$ is the contrast transfer function, which we assume here to be radially symmetrical. $CTF(s)$ can be approximated as $CTF(s) = C(s)E(s)$, where $E(s)$ is a function that dampens real space high resolution terms, and $C(s)$ is an oscillating function with increased periodicity at high resolution [22]. The CTF essentially is a dampened oscillating real function that passes through zero many times.

$F(s)$ is the structure factor assuming the kinematic approximation [22] and $N(s)$ is Fourier transform of the detector readout and quantum noise. Strictly speaking, $F(s)$ has a random component too, caused by disordered (solvent) density. This term is usually ignored, as it is subject to the same corrections as the structure factors corresponding to ordered density. Estimation procedures determine the parameters of the functions $CTF(s)$ and $N(s)$ to optimally fit the observed power spectral curve of rotation average of $M(s)$.

* Corresponding author. Tel.: +31 71 5274211; fax: +31 71 5274357.
E-mail address: L.Jiang@science.leidenuniv.nl (L. Jiang).

0304-3991/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.ultramic.2010.01.011

Different researchers may use different denotations for the frequency variable s, for example f, k, etc. Here we use s uniformly. The detailed formulation of the functions may also differ slightly in the different software packages.

Once estimates of the CTF and noise parameters are available, estimates of the functions of CTF and noise will be known. There are several solutions to use these in correcting the measured image data in 3D reconstruction software packages:

1. Filtering at the first zero of the CTF by truncating the high resolution part after the first zero. No actual CTF correction is applied in this case. Usually it is suggested to use this procedure only for making the first prototype model and in other early stages of the structure determination.
2. Applying phase correction only—as it is, for instance, done in IMAGIC [13]. This is achieved by flipping phases of structure factors at spacings where the CTF dips below zero, whilst keeping the amplitudes intact. The rationale of flipping the phases is that the phase plays a much more important role in the structure determination than the amplitude [23]. The rationale for not correcting the amplitudes is that boosting low level amplitudes close to the CTF zeroes will deteriorate the overall signal-to-noise ratio in rings in Fourier space. Hence, only applying phase correction without bothering about the amplitudes, also has practical advantages.
3. Do both phase and amplitude correction. Complete CTF correction (or full CTF correction) is normally performed in two separate steps, first flipping the phase, and then applying amplitude correction. Due to it being theoretically optimal, the problem of full CTF correction is frequently addressed in the community of 3DEM methods research (e.g. [11,24–29]).

A general approach to do full CTF correction is to find a deconvolution filter function $G(s)$ so that we can estimate $F(s)$ as follows:

$$\hat{F}(s) = G(s)M(s) \tag{2}$$

To recover the amplitude of the object $F(s)$, a simple attempt is

$$\hat{F}(s) = (1/CTF(s))M(s) \tag{3}$$

Here $G(s)=1/CTF(s)$. However, this attempt is not feasible in practice due to the problems of random noise and zeros of the CTF. The random noise cannot be removed directly[1]. It is expected to be reduced by averaging multiple images in one orientation class[2]. The CTF has many zeros with the changing of phase, it is relatively small at low frequencies and tends to zero at the high frequency end due to the shape of the envelope function. The restored image will be corrupted by noise, which will be enhanced upon division by the CTF in regions where the CTF is small [32]. All these features of CTF render the straightforward division by the CTF sub-optimal.

In full CTF correction, after the phase is flipped, several methods may be employed in amplitude correction to avoid dividing by zero and to prevent amplifying the noise while deconvoluting the contrast transfer function:

1.1. Wiener deconvolution

The Wiener filter is used widely in image processing [33]. An application of the Wiener filter [34] is used for amplitude correction (e.g. in SPIDER, EMAN). The Wiener deconvolution filter can be formulated in the frequency domain as follows:

$$G(s) = \frac{1}{H(s)}\left[\frac{|H(s)|^2}{|H(s)|^2 + 1/SNR(s)}\right] \tag{4}$$

Here $H(s)$ is the frequency transfer function, $1/H(s)$ is the inverse of the original system, corresponding to $1/CTF(s)$ in the CTF correction. $SNR(s)=S(s)/N(s)$ is the signal-to-noise ratio (SNR), S(s) is the signal intensity ($=CTF(s)^2F(s)^2$) and N(s) is the noise intensity ($=N(s)^2$).

In order to use the Wiener filter, one has to estimate or determine the SNR. Consequently, solution structure factors (the rotationally averaged curve of $F(s)$) need to be estimated independently, e.g. by a small angle X-ray scattering (SAX) experiment.

When there is low noise (SNR is very large), the term in the square brackets tends to 1, and the Wiener filter equals approximately the inverse of $H(s)$. However, when the noise is strong (SNR is very small), the term in the square brackets will decrease, thus suppressing the intensity of the noise—note that in this case also the signal is suppressed strongly. The term within the square brackets is therefore a kind of amplitude optimization, fine-tuning the amplitude of the restored signal to minimize the mean square error between the original and the estimated signal.

The Wiener filter cannot recover missing information in the zero regions, and an adapted Wiener filter is needed to mediate the information of different defocus images at the same frequency and generate an integrated image. In 3D reconstruction, a set of images assigned to the same orientation class is used in calculating such an integrated image (or class average image). Application of a Wiener filter in 3D reconstruction was described by Penczek et al. [32]. To describe the filter of the n'th data set in a formula (the notation is adapted here for convenience):

$$G(s) = \frac{SNR_n CTF_n^*(s)}{\sum_{n=1}^{N} SNR_n |CTF_n(s)|^2 + 1} \tag{5}$$

where $CTF_n^*(s) = \frac{1}{CTF_n(s)}|CTF_n(s)|^2$. Collecting a defocus series data set covering the whole range of frequencies from zero to some limit of frequency of sampling, the adapted filter combines the data sets and performs CTF correction in Fourier space.

The application of the Wiener filter in 3D reconstruction needs an estimate of the spectral SNR, e.g. an X-ray scattering curve (solution structure factor) is necessary for this purpose. However, this is unavailable in many cases.

Moreover, the assumption that we have a sufficient number of different defocus images and the CTFs can jointly cover the whole Fourier space without a gap is not always true. For instance, in the reconstruction with a small angular sampling step for projections (e.g. 3 degrees), more than one thousand projections/classes can be used (especially for a model of C1 symmetry); lots of classes contain a few particles only (e.g. less than 10) as a basis for generating a class average image. The Wiener filter method is not optimal in this case due to the large probability of superposition of multiple zeros. An accurate estimate of the CTF parameters is essential, otherwise the merging of information pertaining to different particle images at the same frequency will lead to a breakdown of the continuity of the image in Fourier space.

[1] We do not discuss approaches that reduce noise by improved detectors or other experimental aspects of data collection (Medipix: a photon counting pixel detector [30,31]), as these approaches are fully compatible with the improvements in data analysis discussed here.

[2] With orientation class we mean the result of references/projections supervised classification or an automatic classification. In an orientation class, images are assumed to be the projections from the same view of a 3D model and they are used to calculate a class average image.

1.2. Spatial frequency weighted averaging

Performing a weighted average of the images, where the weights vary with spatial frequency. Ludtke et al. [11,35] use weight factors to avoid dividing by zero in amplitude correction. The weight factors in averaging the images in one orientation class ($K_n(s)$) are given by

$$K_n(s)=\frac{1}{C_n(s)E_n(s)}\frac{R_n(s)}{\sum_m R_m(s)}=\frac{C_n(s)E_n(s)}{\sum_m C_m(s)^2E_m(s)^2} \quad (6)$$

where the subscript 'n' denotes particle number and 'm' denotes the total amount of particles in the orientation class. The term $1/(C(s)E(s))$ is the inverse of CTF (the same function as the term $1/H(s)$ of the Wiener filter). $R_n(s)=C_n(s)^2E_n(s)^2/N_n(s)^2$ is used as the relative signal-to-noise ratio (SNR) for each particle. $N_n(s)^2$ is left out, assuming it to be approximately equal in different micrographs. If an estimate of the solution structure factor curve is known, the absolute SNR can be calculated and used instead of the relative SNR. In this case, this method actually acts as a Wiener filter. An additional Wiener filter or a low-pass Gaussian filter may still be applied to smooth the final model.

If there are only a few defocused images, this procedure may run into trouble of coincident zeros and in practice EMAN calculates the direct average.

1.3. Other methods.

Other ways of doing a full CTF correction have been tried as well, such as the iterative method given by Penczek et al. [32], the iterative data refinement (IDR) technique [25], and the Chahine's method [27]. Differing in important details, these methods all attempt finding an approximation of the original image by iterative refinement or minimization of a residual function. Their penalty is that they increase the time consumed by the 3D reconstruction.

CTF correction is a vital prerequisite for effective 3D reconstruction using 2D images obtained by electron microscopy. This explains why so many researchers are continuously, trying to improve existing methods and developing new ones. Here, we introduce a novel approximation filter for CTF correction. It is easily implemented and shows good convergence properties in iterative reconstruction refinement. Application of the filter proposed clearly improves the resolution and it is robust in tests with noisy, close-to-focus data sets.

2. Method

We propose a novel approach: we constructed continuous and differentiable function, which allows direct application of an inverse CTF filter in 3D reconstruction. The function contains no singularities in its approximation to the CTF. Since the simple attempt of $G(s)=1/CTF(s)$ fails because of $CTF(s)$ having zeros and other small values, the CTF curve is partially modified, avoiding zeros and preventing divisions by small values at high spatial frequencies. The modified curve must be continuous to avoid edge effects in Fourier filtering.

The reasons for trying this new approach are

- There is no need to separately estimate the true structure factor (without the noise), so it can even be applied to a single image.
- If the CTF is not known very accurately, the method includes integration over the uncertainties of the CTF (see Fig. S1).
- Differentiable functions in general produce fewer artifacts in filtering than functions with sharp edges, as sharp edged functions cause ripples upon Fourier transformation.
- Continuous, differentiable functions also can be minimized much more effectively, as first (and second) derivative information is very valuable in determining the search direction in refinement procedures [36].

The proposed inverse filter can be described as

$$G(s)=\begin{cases}1/(C(s)\hat{E}(s)) & C(s)>=0.5\\ 1/((1-\sqrt{0.5-C(s)^2})\hat{E}(s)) & C(s)<0.5\end{cases} \quad (7)$$

$$\hat{E}(s)=E(s)Sig(s-s_0)+N(s)\alpha(1-Sig(s-s_0)) \quad (8)$$

Here, $\hat{E}(s)$ is an estimation of $E(s)$, $Sig(s-s_0)$ is a sigmoid function ($Sig(x)=1/(1+e^{-x})$) as is shown in Fig. 1. The idea is to create a continuous and differentiable function (also continuous for derivatives of higher order) to 'glue' together the functions of $E(s)$ at low frequency region and $N(s)$ at the high frequency region, where the noise dominates the density measured. The scale factor α scales $N(s)$ to the same level of $E(s)$ at joint point S_0. The user selected value S_0 defines the frequency joint point of $N(s)$ and $E(s)$. The sensible choice is $S_0=1/\sqrt{B}$, where B is the envelope B-factor of the CTF estimate. At this point, the SNR value diminishes quickly.

The modified $G(s)$ is a piecewise continuous function, which is continuous also in its first derivative. At the region near zeros (where $C(s)<0.5$), the $C(s)$ is modified to a piece of continuous arc $(1-\sqrt{0.5-C(s)^2})$, which has a minimum value of approximately 0.2929 at the original zeros. The inverse of 0.2929 is a small number ($\sim$3.4). This simple modification makes the inverse deconvolution method feasible (Fig. 2 shows the curve of the approximation to the CTF, the numerator of $G(s)$). At the high frequency region, the numerator of $G(s)$ still tends to zero, however, the estimation of $E(s)$ ($\hat{E}(s)$) tends to be of the same order as $N(s)$. After filtering by $G(s)$, the intensity of the signal and the noise at high frequencies are only multiplied by a small number.

The proposed function of $G(s)$ is an approximation assuming noise and small uncertainties in the CTF parameters. In the absence of noise and full knowledge of the CTF, a better function can be formulated.

In the 3D reconstruction, noise is decreased in two stages of averaging: first in generating each of the class average images, then in 3D reconstruction, when thousands of class average images are combined to form a 3D model. By applying the new inverse filter the noise is amplified somewhat for a single particle image, but by a limited factor only (a maximum around 3.4 times at zeros in low frequency region). Despite the new filter

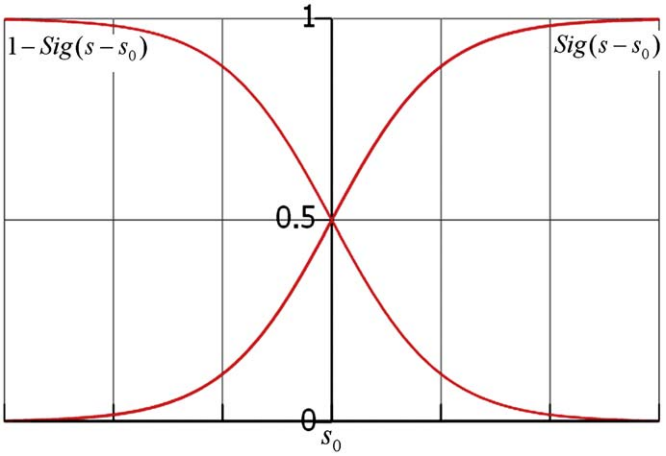

Fig. 1. The functions $Sig(s-s_0)$ and $1-Sig(s-s_0)$

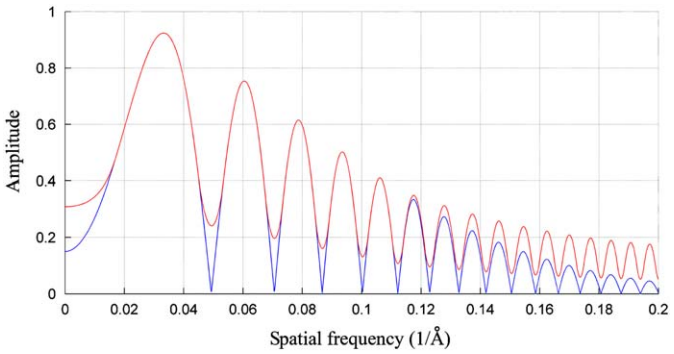

Fig. 2. Blue: Theoretical CTF curves after phase flipping. Red: The approximation of the CTF (=1/*G(s)*) used in the inverse filter (after phase flipping) (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

increasing both the signal and noise near zeros of the CTF, we show the averaging procedure to reduce the noise efficiently: the noise is still under control and the algorithm is stable.

Two new algorithms that apply the proposed function of *G(s)* have been implemented for calculating a class average image.

Algorithm of "direct CTF deconvolution": A class average image can be calculated by aligning and straightforward averaging of all the *G(s)* filtered images within an orientation class. To distinguish our method from other conventional methods, we refer to it as "direct CTF deconvolution", as its filter *G(s)* can be used for calculating the class average image without a Wiener filter, which in a formula is $\hat{F}_c(s) = \frac{1}{N}\sum_1^N G(s)M(s)$. A wide low-pass Gaussian filter can still be used to erase the limited noise at high resolution. A Wiener filter or other filter can be an optional choice for the later processing stages, but by omitting it at this early stage, we can also circumvent some of its more problematic aspects (see above).

Algorithm of "filtered CTF correction": The proposed function of *G(s)* can also be used in combination with other full CTF correction methods, since most of them have an inverse term in the algorithm. For example, in the weighted average method (which is used by the 'ctfc' or 'ctfcw' options of EMAN software), the new function of *G(s)* can be used instead of the first term $1/(C(s)E(s))$ of the weight factors in averaging the images. In this case, the implemented algorithm works like a selective filter, suppressing the noise only at regions near zeros and the high resolution end, in a manner adapted to each image.

These two new algorithms were implemented and tested in EMAN 1.6, and it also works for the newer version of EMAN 1.8. The algorithm of "filtered CTF correction" is implemented as a combination of the new function with the weighted average method of the EMAN refinement program, which also applies a Wiener filter.

We have tested and compared the new algorithms using a data set from a highly symmetric structure and one obtained from a structure with no internal symmetry.

3. Results

Two sample data were used to test the algorithms. First, we demonstrate the feasibility of the two new algorithms. Then, by comparing the results from conventional full CTF correction to the new algorithms, we show that the new algorithms both have the advantage of better convergence, while both are stable in the reconstruction process of asymmetric macromolecules.

3.1. Highly symmetrical particles: a test with GroEL

The high resolution EM data of native free GroEL that we used, were kindly provided by Ludtke and Chiu for testing the algorithms. The data were first made available course material for the participants of the workshop on Single Particle Reconstruction and Visualization 2007 in Houston, USA. Sample preparation and data acquisition were described elsewhere [37]. GroEL is a homotetradecameric protein consisting of two back-to-back stacked rings, each of them containing seven identical subunits. The rings have an outer diameter of 13.7 nm and inner diameter of 4.5 nm; each monomer has a molecular weight of 58 kDa [38,39]. GroEL is one of the typical test-beds for methods research in the field of 3DEM reconstruction [40,41]. A total 4169 particles (128×128 pixels in size, 2.08 Å/pixel) were selected from 12 micrographs with defocus ranging from 1.9–2.3 μm. D7 symmetry (14 fold symmetry) is imposed in the reconstruction, thus 58366 asymmetric subunits are used (Fig. 3).

The algorithms were all tested using the same particle data set and starting model. The same set of projections was used as class references. The classification required for making the averaged images was imported from one single reference-supervised-classification procedure. Only the algorithms for making the average image (including full CTF correction) were different. The three models shown in Fig. 4 were reconstructed from the same reference model and classification.

In Fig. 4, the result of "the direct CTF deconvolution" algorithm (middle model, M2) shows more detail than the conventional CTF correction algorithm (left model, M1). The result of "the filtered CTF correction" algorithm (right model, M3) looks like a low-pass-filtered model of the normal CTF correction. When the approximation deconvolution filter is combined with the normal CTF correction, it works more like a selective filter, suppressing the noise strongly.

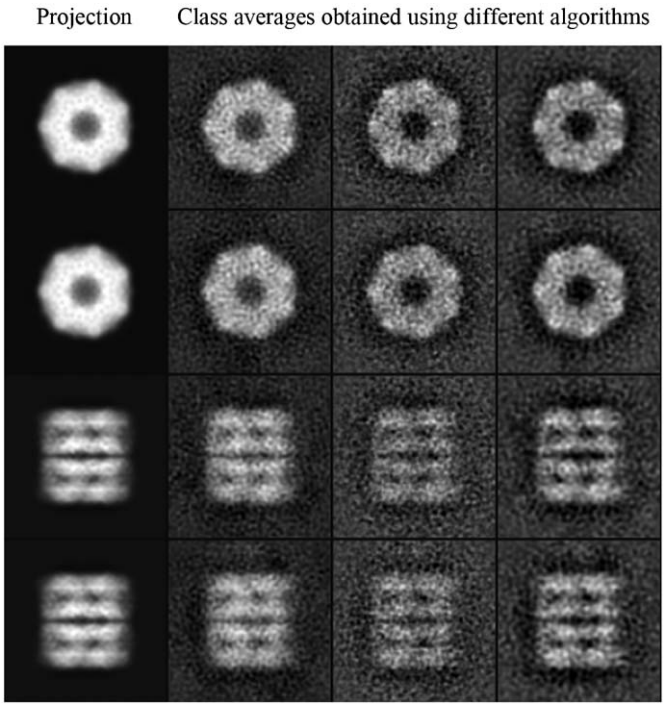

Fig. 3. Representative class averages of GroEL generated by different CTF correction algorithms. The second column is generated by the conventional full CTF correction algorithm. The third column is generated by the new direct CTF deconvolution algorithm. The last column is generated by the new filtered CTF correction algorithm, which combines the approximated inverse filter and the weighted average method.

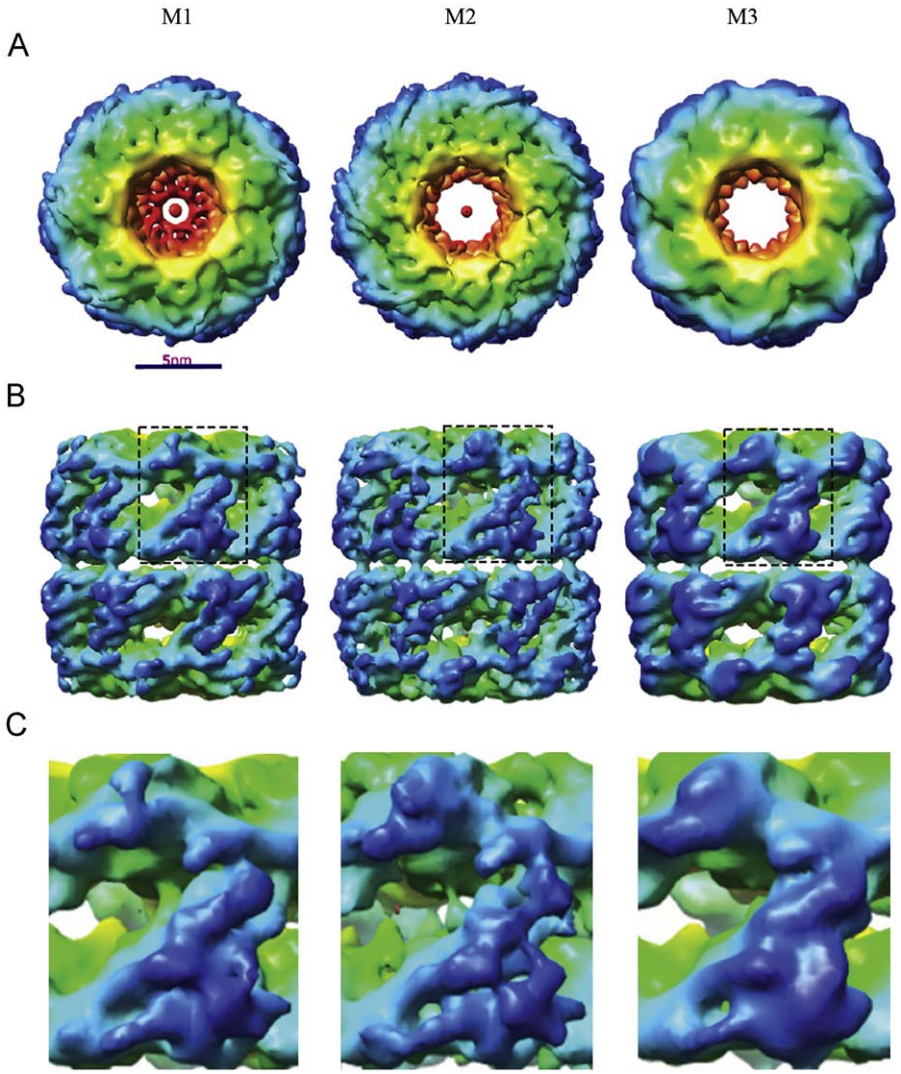

Fig. 4. Reconstructed models of GroEL generated with different CTF correction algorithms. M1(column left): The model obtained by conventional full CTF correction, resolution 7.9 Å; M2(middle): The model obtained by the "direct CTF deconvolution" algorithm, 7.0 Å; M3(right): The result of the "filtered CTF correction" algorithm, 8.6 Å. (A) Top views of three different models. (B) Side views corresponding to A. The black dash line boxes indicate subunits for visually comparison. (C) Views zoomed in for the indicated subunits in B.

The result of "the direct CTF deconvolution" algorithm–M2 has the best resolution based on an even-odd test and Fourier shell correlation (FSC [42]) with a threshold of 0.5 [43]: 7.0 Å (Table 1). It should be realized that such a test does not measure the overall resolution, but is a reflection of the internal consistency between model and data. This can be clouded by bias introduced by specific choices of envelopes, samplings, symmetries and starting models [44], and we were careful to ensure that these were identical for the various algorithms we tested. The FSC at 0.5 of M2 is around 1 Å better than that of M1, while the M3 is a little bit over-filtered and lost some high resolution detail. We stress that the FSC criterion cannot be used to compare the resolution of GroEL with that of the ribosomal complex mentioned below, as there were differences in symmetry, box size, sampling, etc. The electron density (Figs. 4 and 6) does indicate the resolution of the GroEL reconstruction to be considerably higher.

One will also notice that the densities in the inner channels of these three models are apparently different. It seems that M2 and M3 have less resolved structure in the channel, but actually there is almost no density in the inner channel (within 4.5 nm diameter) of the X-ray structure of unliganded GroEL (PDBid: 1OEL, [45]). A recent published higher resolution (~4 Å) result of GroEL using cryo-EM reconstruction [41] also excludes the possible existence of additional ordered density in the inner channel of "live" GroEL in solution. All this confirms the validity of the new CTF correction algorithms. We will further address the "correctness" of the EM models in the discussion.

In order to compare the convergence of different algorithms, we also calculated the Fourier shell correlation curves between the results and the starting model for a rough comparison of the inter-models differences (Table 1). Here the FSC value between models can be considered as a measure of inter-model similarity. From the comparison, we can see that M2 is 6.7 Å similar to the starting model, while M1 is at 5.3 Å. It shows that the "direct CTF deconvolution" method has better convergence properties: less model bias (it is more different from the starting model) and smaller internal divergence (better resolution, more consistency between the models of even and odd numbered particles).

The test shows the two new algorithms to be feasible alternatives for other CTF correction algorithms, and suggests that they have better convergence properties than the normal full CTF correction.

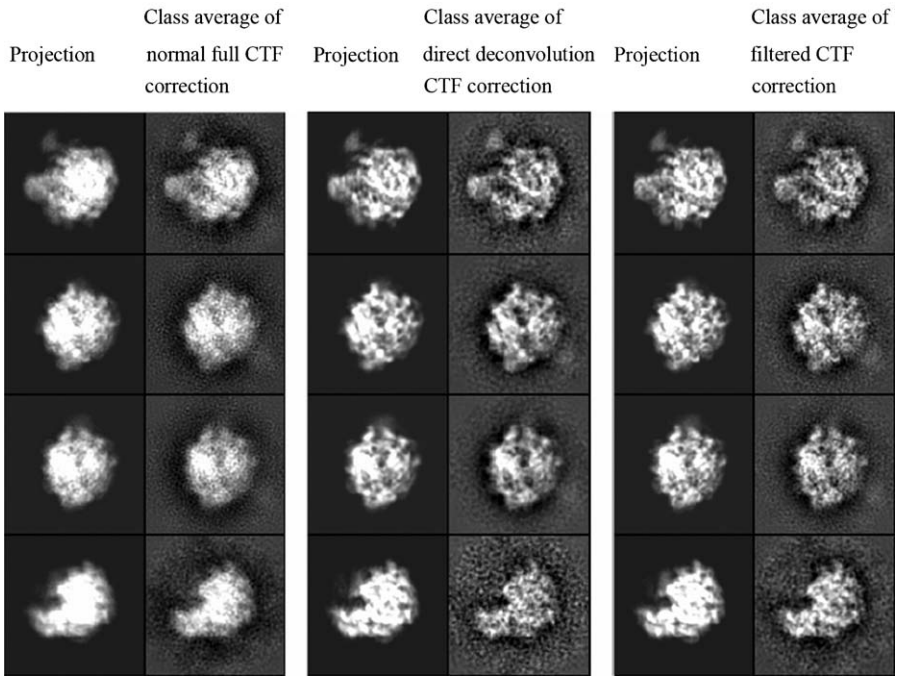

Fig. 5. Representative class average images of the complex of ribosomal 50S particle with nascent chain tRNA and Hsp15. The projections and average images are selected from the fourth round of iterative refinement (well before the conventional CTF correction algorithm blew up). Average images in the left panel are generated with the conventional full CTF correction. Average images in the middle panel are generated with "direct CTF deconvolution" algorithm. Average images in the right panel are generated with "filtered CTF correction" algorithm.

Table 1
Comparison of resolutions and inter-models similarity.

Model	M1	M2	M3
Resolution (FSC[a] at 0.5) (Å)	7.9	7.0	8.6
Similarity with starting model (FSC at 0.5) (Å)	5.3	6.7	8.3

Fourier shell correlation (FSC, 0.5 criterion) of different models of GroEL (M1: conventional CTF correction; M2: direct CTF convolution; M3: filtered CTF correction). The table indicates that for highly symmetrical particles the direct CTF convolution algorithm converges to a more self-consistent model that suffers less from model bias.

[a] FSC, Fourier shell correlation.

3.2. Asymmetrical particles: the stalled ribosomal 50S complex

The complex of a large ribosome subunit 50S with tRNA and heat shock protein 15 was used to test the new algorithms. The complex has a diameter of ~20 nm and weight of ~1,600 kDa. We used a set of 33,900 images containing 128×128 pixels recorded with a defocus ranging from 0.6 to 1.8 μm [46]. We collected close-to-focus micrographs with a low-dose exposure ($< 10e^{-}/Å^2$), and therefore collected relatively noisy images. Single particles were selected and maintained by using the Cyclops software [47]. For the 981 projections/classes used in the reconstruction, the average number of images per orientation class is around 34.

A model that had nearly converged to the final model (from [46]), was used as a common starting model for all the algorithms tested. After four iterations of refinement using the same data set and starting model, but using different CTF correction algorithms, we compared the class average images and 3D reconstruction results in Figs. 5 and 6. If we ran more rounds of iterative refinement, the method of normal CTF correction blew up because of the high level of noise in the test data.

Fig. 5 shows some selected, representative average images from orientation classes having relatively many particles. Comparison of the projections and class average images with the results obtained from the GroEL samples in Fig. 3 shows the ribosome projections to be more blurred and have less contrast. The reasons are mainly:

- The GroEL data set of 4169 particles contains the highest contrast particles selected from the original 39,085 particles. Next they were used to reconstruct a 6 Å structure [37]. In contrast, the data of the ribosome complex were collected with a relatively smaller defocus (near 1 μm), resulting in a lower contrast and more noisy images. Better contrast of the original images used in reconstruction usually results in better contrast in the average images.
- The effective number of particles used in reconstruction for the GroEL is much larger than the number used for the ribosomal subunit due to the high symmetry of GroEL, causing the SNR to be higher in the average images of GroEL.
- The asymmetric ribosome structure is more irregular, thus the projections looks less like stripes and more like blobs.

In Fig. 5, the averaged images of the new algorithms (images in the middle and right columns) look better with more detail than the conventional results (images in left column). The substantial improvement in the quality of the class average images directly results in better 3D models being reconstructed, as is shown in Fig. 6.

In Fig. 6, the results of "filtered CTF correction" (right model) and "direct CTF deconvolution" (middle model) show similar fine detail. Many high resolution structural features of rRNA's double helices can be observed including, for instance, the turns and major and minor grooves. Even the general shapes and densities of helices of the ribosomal proteins can be recognized, while the result of normal full CTF correction (left model) is generally over-corrected.

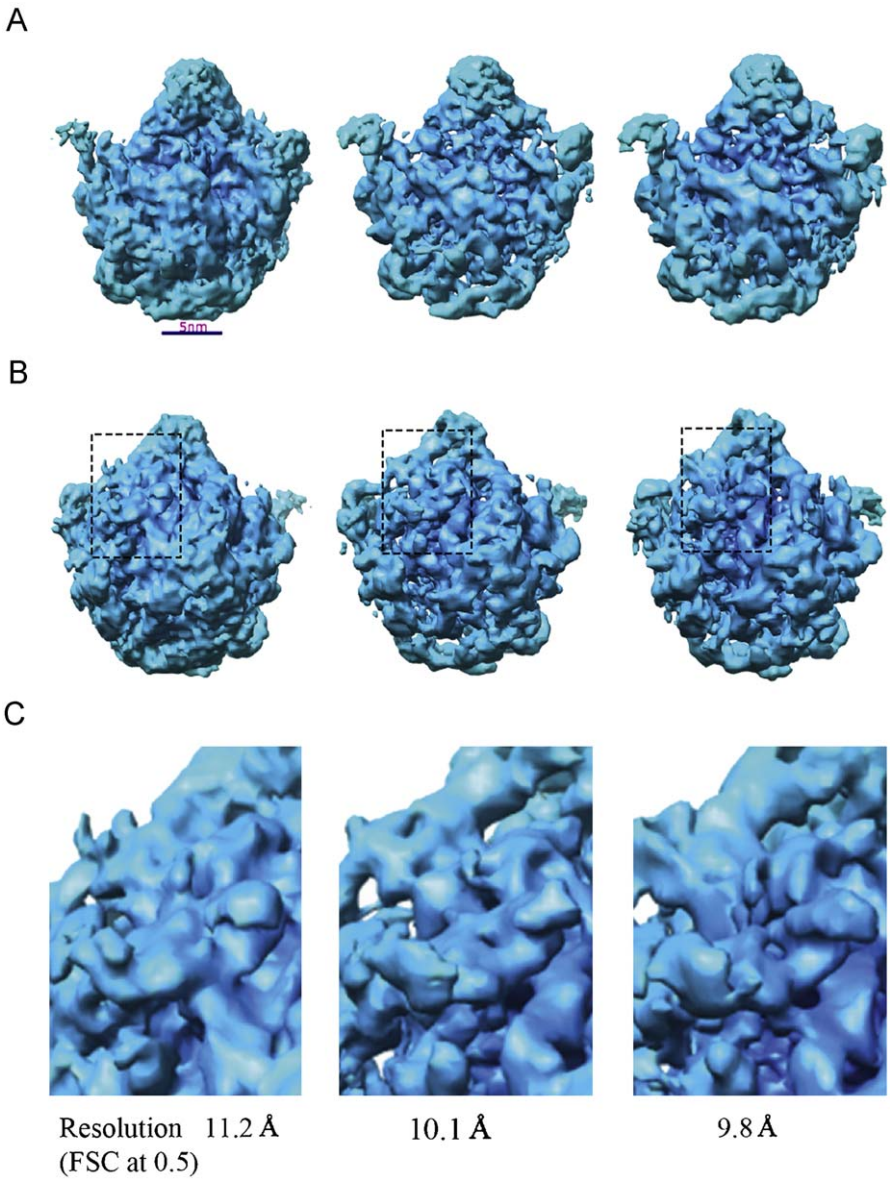

Fig. 6. 3D models reconstructed with different CTF correction algorithms. Column left: The result of conventional full CTF correction as implemented in most used reconstruction software after four iterations of refinement from a starting model. Middle and Right: The results of "direct CTF deconvolution" and "filtered CTF correction" algorithms after four iterations of refinement using the same starting model and the same data set as used for the left model. (A) Front view of ribosomal 50S complex. (B) Back view of ribosomal 50S complex. The black dash line boxes indicate representative regions. (C) Views zoomed in for the indicated regions in B.

Using even-odd tests and FSC with 0.5 criterion, we calculated the resolution of each model. The model obtained using the "direct CTF deconvolution" has a resolution of 10.1 Å. The model of "filtered CTF correction" has an even slightly better resolution of 9.8 Å. In both cases, the resolution improved by more than 1 Å compared to the conventional full CTF correction, using the same data and starting model. The differences between the results on the ribosome compared to the results of GroEL suggest that it is better to use the "filtered CTF correction" algorithm when the experimental data are very noisy (e.g. images with a defocus of less than 1 μm). Although the data are noisy, the algorithm stably suppresses the noise. The "filtered CTF correction" algorithm suppresses the noise more strongly, but does not prohibit the 3D reconstruction. On the contrary. Using only four iterative cycles, the new algorithms are capable of producing a 3D structure with a better resolution, while converging to the optimal resolution more quickly.

4. Discussion

The "direct CTF deconvolution" algorithm provides a novel approach to full CTF correction without using a Wiener filter. Although it is an approximation approach, in our tests it gave better results in practice than current methods. Especially when we have enough particle projections or the particles are highly symmetric, the new algorithms are probably very useful.

Our method slightly over-filters at high resolution (see Fig. 2, the red curve is higher than the blue curve at high resolution, resulting in a dampening of these frequencies upon correction). One should be careful with procedures that boost high spatial frequencies, as they may lead to over-fitting.

In order to push the resolution of the final model in 3DEM reconstruction beyond 1 nm, the user normally needs to do many iterations until the refinement has converged. This refinement step is important and represents the most time consuming step in

the reconstruction procedure. Since refinement does not always converge, the new algorithms may provide at least part of the answer to this problem.

One reason for failure of the iterative refinement is distance in conformational space between the starting reference model and the final model. The starting model normally has a resolution lower than 2 nm, the expected final model has a resolution higher than 1 nm. It would be very helpful and important to get stable intermediate-resolution (between 1 and 2 nm) models in the reconstruction, filling the gap between the starting model and the final high resolution model, leading the iterative reconstruction to converge to the correct density map. For these stable intermediate-resolution models, it is more important to be "correct" rather than having lots of uncertain detail, in order to avoid model bias when we are trying to push the final model towards atomic resolution.

The new filter is efficient in getting a stable intermediate-resolution model, in combination with the existing full CTF correction method (implemented and tested as the "filtered CTF correction" algorithm). In this combination, the algorithm also operates as an extra filter suppressing the noise more strongly. Although the result may have a little bit less resolution, the enhancement of the signal-to-noise ratio improves the stability of the model. For instance, in the first test of GroEL, this combined method did not generate the highest resolution model, but it did produce a stable intermediate-resolution model: the observation of less (spurious) density in the inner channel of M3 is in line with the X-ray structure of GroEL and the higher resolution EM model, as presented in the results section. It proves that such stable model is also a "correct" model.

How much of the improvement in the model resolution is due to better alignment, and how much is due to the new CTF correction method? We cannot answer this issue unequivocally. The alignment and the CTF correction are so closely linked (a better CTF correction allows a better alignment) that we cannot separate the two. Only the first iteration, before alignment, therefore could settle this issue. The differences are so small in this case that we assume the major improvement of our method causes from improved alignment due to the alternative CTF correction.

5. Conclusion

A new approximation method of CTF correction has been implemented which was implemented in the EMAN package and can straightforwardly be incorporated into other 3D reconstruction software packages. It provides an alternative CTF correction method in generating class average images. The method shows better convergence and a better resolution of the final 3D structure, even in the presence of relatively high noise levels. The applied filter is continuously differentiable and does not introduce Fourier artifacts. It effectively avoids instability in regions where the CTF has zeros. The CTF correction is thus less sensitive to the zeros of CTF. When the zeros of CTF are not or cannot be estimated accurately, e.g. due to the low contrast, slightly drift or astigmatism of the micrographs, the new algorithm is expected to have higher tolerance.

The approximation inverse filter is only partially modified from the inverse function. When calculating the class average image, images with different defocus will compensate each other at the zeros of the CTF in the Fourier space. It must converge to the complete CTF correction when the number of differently focused images that make up the average images increases.

Acknowledgements

We thank Wah Chiu and Steven Ludtke for the permission of using the GroEL images data. We are grateful to Dr. RAG de Graaff for revising the manuscript. The figures for 3D models were made by using UCSF's Chimera [48]. This work was financially supported by the Cyttron Foundation [49].

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.ultramic.2010.01.011.

References

[1] Z.H. Zhou, Towards atomic resolution structural determination by single-particle cryo-electron microscopy, Current Opinion in Structural Biology 18 (2008) 218–228.
[2] W. Chiu, M.L. Baker, W. Jiang, M. Dougherty, M.F. Schmid, Electron cryomicroscopy of biological machines at subnanometer resolution, Structure 13 (2005) 363–372.
[3] F. Thon, Zur Defokussierungsabhangigkeit des Phasenkontrastes Bei der Elektronenmikroskopischen Abbildung, Zeitschrift fur Naturforschung Part A-Astrophysik Physik und Physikalische Chemie A 21 (1966) 476–478.
[4] H.P. Erickson, A. Klug, Fourier transform of an electron micrograph—effects of defocussing and aberrations, and implications for use of underfocus contrast enhancement, Berichte der Bunsen-Gesellschaft fur Physikalische Chemie 74 (1970) 1129–1137.
[5] M. van Heel, W. Keegstra, Imagic—a fast, flexible and friendly image–analysis software system, Ultramicroscopy 7 (1981) 113–130.
[6] M. van Heel, G. Harauz, E.V. Orlova, R. Schmidt, M. Schatz, A new generation of the IMAGIC image processing system, Journal of Structural Biology 116 (1996) 17–24.
[7] J. Frank, B. Shimkin, H. Dowse, Spider—a modular software system for electron image-processing, Ultramicroscopy 6 (1981) 343–357.
[8] J. Frank, M. Radermacher, P. Penczek, J. Zhu, Y.H. Li, M. Ladjadj, A. Leith, SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields, Journal of Structural Biology 116 (1996) 190–199.
[9] C.O.S. Sorzano, R. Marabini, J. Velazquez-Muriel, J.R. Bilbao-Castro, S.H.W. Scheres, J.M. Carazo, A. Pascual-Montano, XMIPP: a new generation of an open-source image processing package for electron microscopy, Journal of Structural Biology 148 (2004) 194–204.
[10] R. Marabini, I.M. Masegosa, M.C. SanMartin, S. Marco, J.J. Fernandez, L.G. delaFraga, C. Vaquerizo, J.M. Carazo, Xmipp: an image processing package for electron microscopy, Journal of Structural Biology 116 (1996) 237–240.
[11] S.J. Ludtke, P.R. Baldwin, W. Chiu, EMAN: semiautomated software for high-resolution single-particle reconstructions, Journal of Structural Biology 128 (1999) 82–97.
[12] Y.Y. Liang, E.Y. Ke, Z.H. Zhou, IMIRS: a high-resolution 3D reconstruction package integrated with a relational image database, Journal of Structural Biology 137 (2002) II.
[13] M. van Heel, B. Gowen, R. Matadeen, E.V. Orlova, R. Finn, T. Pape, D. Cohen, H. Stark, R. Schmidt, M. Schatz, A. Patwardhan, Single-particle electron cryo-microscopy: towards atomic resolution, Quarterly Reviews of Biophysics 33 (2000) 307–369.
[14] Z.H. Zhou, S. Hardt, B. Wang, M.B. Sherman, J. Jakana, W. Chiu, CTF determination of images of ice-embedded single particles using a graphics interface, Journal of Structural Biology 116 (1996) 216–222.
[15] Z. Huang, P.R. Baldwin, S. Mullapudi, P.A. Penczek, Automated determination of parameters describing power spectra of micrograph images in electron microscopy, Journal of Structural Biology 144 (2003) 79–94.
[16] J.J. Fernandez, S. Li, R.A. Crowther, CTF determination and correction in electron cryotomography, Ultramicroscopy 106 (2006) 587–596.
[17] J.A. Velazquez-Muriel, C.O.S. Sorzano, J.J. Fernandez, J.M. Carazo, A method for estimating the CTF in electron microscopy based on ARMA models and parameter adjustment, Ultramicroscopy 96 (2003) II.
[18] S.P. Mallick, B. Carragher, C.S. Potter, D.J. Kriegman, ACE: automated CTF estimation, Ultramicroscopy 104 (2005) 8–29.
[19] B. Sander, M.M. Golas, H. Stark, Automatic CTF correction for single particles based upon multivariate statistical analysis of individual power spectra, Journal of Structural Biology 142 (2003) 392–401.
[20] F. Thon, in: U. Valdre (Ed.), Phase Contrast Electron Microscopy, Electron Microscopy in Materials Science Academic Press, New York, 1971, pp. 570–625.
[21] K.J. Hanszen, The optical transfer theory of the electron microscope: fundamental principles and applications, in: R. Barer, V.E. Cosslett (Eds.), Advances in Optical and Electron Microscopy, , 1971, pp. 1–84.

[22] J. Frank, in: Three-Dimensional Electron Microscopy of Macromolecular Assemblies, Academic Press, San Diego, 1996.
[23] G.N. Ramachandran, R. Srinivasan, in: Fourier Methods in Crystallography, Wiley, New York, 1970, pp. 60–71.
[24] N. Grigorieff, FREALIGN: high-resolution refinement of single particle structures, Journal of Structural Biology 157 (2007) 117–125.
[25] C.O.S. Sorzano, R. Marabini, G.T. Herman, Y. Censor, J.M. Carazo, Transfer function restoration in 3D electron microscopy via iterative data refinement, Physics in Medicine and Biology 49 (2004) DOI.
[26] Wah Chiu Yi Wan, Z.H. Zhou, Full contrast transfer function correction in 3D cryo-EM reconstruction, International Conference on Communications, Circuits and Systems 2 (2004) 960–964.
[27] J.P. Zubelli, R. Marabini, C.O.S. Sorzano, G.T. Herman, Three-dimensional reconstruction by Chahine's method from electron microscopic projections corrupted by instrumental aberrations, Inverse Problems 19 (2003) II.
[28] J. Zhu, P.A. Penczek, R. Schroder, J. Frank, Three-dimensional reconstruction with contrast transfer function correction from energy-filtered cryoelectron micrographs: procedure and application to the 70S *Escherichia coli* ribosome, Journal of Structural Biology 118 (1997) 197–219.
[29] J. Frank, P. Penczek, On the correction of the contrast transfer-function in biological electron-microscopy, Optik 98 (1995) 125–129.
[30] Medipix, 〈http://medipix.web.cern.ch/MEDIPIX/〉.
[31] J.R. Plaisier, R.I. Koning, H.K. Koerten, A.M. van Roon, E.A.J. Thomassen, M.E. Kuil, J. Hendrix, C. Broennimann, N.S. Pannu, J.P. Abrahams, Area detectors in structural biology, Nuclear Instruments & Methods in Physics Research Section A-Accelerators Spectrometers Detectors and Associated Equipment 509 (2003) 274–282.
[32] P.A. Penczek, J. Zhu, R. Schroder, J. Frank, Three dimensional reconstruction with contrast transfer compensation from defocus series, Scanning Microscopy (1997) 147–154.
[33] R. Gonzalez, R. Woods, S. Eddins, in: Digital Image Processing Using Matlab, Prentice Hall, 2003.
[34] P. Schiske, Image processing using additional statistical information about the object, in: P. Hawkes (Ed.), Image Processing and Computer-aided Design in Electron Optics, Academic Press, London, 1973, pp. 82–90.
[35] S.J. Ludtke, J. Jakana, J.L. Song, D.T. Chuang, W. Chiu, A 11.5 Angstrom single particle reconstruction of GroEL using EMAN, Journal of Molecular Biology 314 (2001) 253–262.
[36] W.H. Press, S.A. Teukolsky, W.T. Vettering, B.P. Flannery, Minimization or maximization of functions, in: W.H. Press (Ed.), Numerical Recipes: The Art of Scientific Computing, Cambridge University Press, 2007.
[37] S.J. Ludtke, D.H. Chen, J.L. Song, D.T. Chuang, W. Chiu, Seeing GroEL at 6 Angstrom resolution by single particle electron cryornicroscopy, Structure 12 (2004) 1129–1136.
[38] F.U. Hartl, Molecular chaperones in cellular protein folding, Nature 381 (1996) 571–580.
[39] K. Braig, Z. Otwinowski, R. Hegde, D.C. Boisvert, A. Joachimiak, A.L. Horwich, P.B. Sigler, The crystal-structure of the bacterial chaperonin GroEL at 2.8-Angstrom, Nature 371 (1994) 578–586.
[40] S.M. Stagg, G.C. Lander, J. Quispe, N.R. Voss, A. Cheng, H. Bradlow, S. Bradlow, B. Carragher, C.S. Potter, A test-bed for optimizing high-resolution single particle reconstructions, Journal of Structural Biology 163 (2008) 29–39.
[41] S.J. Ludtke, M.L. Baker, D.H. Chen, J.L. Song, D.T. Chuang, W. Chiu, De novo backbone trace of GroEL from single particle electron cryomicroscopy, Structure 16 (2008) 441–448.
[42] G. Harauz, M. van Heel, Exact filters for general geometry three dimensional reconstruction, Optik 73 (1986) 146–156.
[43] B. Bottcher, S.A. Wynne, R.A. Crowther, Determination of the fold of the core protein of hepatitis B virus by electron cryomicroscopy, Nature 386 (1997) 88–91.
[44] M. van Heel, M. Schatz, Fourier shell correlation threshold criteria, Journal of Structural Biology 151 (2005) 250–262.
[45] K. Braig, P.D. Adams, A.T. Brunger, Conformational variability in the refined structure of the chaperonin GroEL at 2.8 Angstrom resolution, Nature Structural Biology 2 (1995) 1083–1094.
[46] L. Jiang, C. Schaffitzel, R. Bingel-Erlenmeyer, N. Ban, P. Korber, R.I. Koning, D.C. de Geus, J.R. Plaisier, J.P. Abrahams, Recycling of aborted ribosomal 50S subunit-nascent chain-tRNA complexes by the heat shock protein Hsp15, Journal of Molecular Biology 386 (2009) 1357–1367.
[47] J.R. Plaisier, L. Jiang, J.P. Abrahams, Cyclops: new modular software suite for cryo-EM, Journal of Structural Biology 157 (2007) 19–27.
[48] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, UCSF chimera—a visualization system for exploratory research and analysis, Journal of Computational Chemistry 25 (2004) 1605–1612.
[49] Cyttron, 〈http://www.cyttron.org/〉.