Published in final edited form as:

J Struct Biol. 2009 December ; 168(3): 378–387. doi:10.1016/j.jsb.2009.08.016.

CTF Determination and Correction for Low Dose Tomographic Tilt Series

Quanren Xiong, Mary K. Morphew, Cindi L. Schwartz, Andreas H. Hoenger, and David N. Mastronarde

Boulder Laboratory for Three-dimensional Electron Microscopy of Cells, Department of Molecular, Cellular, and Developmental Biology, University of Colorado, Boulder, CO, 80309

Abstract

The resolution of cryo-electron tomography can be limited by the first zero of the microscope's contrast transfer function (CTF). To achieve higher resolution, it is critical to determine the CTF and correct its phase inversions. However, the extremely low signal-to-noise ratio (SNR) and the defocus gradient in the projections of tilted specimens make this process challenging. Two programs, CTFPLOTTER and CTFPHASEFLIP, have been developed to address these issues. CTFPLOTTER obtains a 1D power spectrum by periodogram averaging and rotational averaging and it estimates the noise background with a novel approach, which uses images taken with no specimen. The background-subtracted 1D power spectra from image regions at different defocus values are then shifted to align their first zeros and averaged together. This averaging improves the SNR sufficiently that it becomes possible to determine the defocus for subsets of the tilt series rather than just the entire series. CTFPHASEFLIP corrects images line-by-line by inverting phases appropriately in thin strips of the image at nearly constant defocus. CTF correction by these methods is shown to improve the resolution of aligned, averaged particles extracted from tomograms. However, some restoration of Fourier amplitudes at high frequencies is important for seeing the benefits from CTF correction.

Keywords

electron tomography; contrast transfer function; cryomicroscopy; software

1. Introduction

The observed power spectrum of projections in electron microscopy is the modulation of the true power spectrum by the (amplitude and phase) contrast transfer function (CTF) that is a result of microscope lens defocus, astigmatism, and aberration, as well as electron wave length and both temporal and spatial coherence of the electron beam [1]. In the images used for three-dimensional reconstruction by cryo-electron tomography (cryo-ET), the CTF is spatially variant because the defocus varies across the tilted specimen in a direction perpendicular to the tilt axis. Furthermore, micrographs of vitrified specimens are typically taken with a certain amount of underfocus to enhance phase contrast, which is the main contrast component in these

© 2009 Elsevier Inc. All rights reserved.

Correspondence to: David N. Mastronarde, Dept. of MCD Biology, University of Colorado, 347 UCB, Boulder, CO 80309, mast@colorado.edu, Phone: (303) 492-4350, Fax: (303) 492-7744.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

images; amplitude contrast contributes only 7 to 14 percent [2]. The resulting combined CTF is a sinusoid-like function that attenuates with increases in spatial frequency.

Precise determination of the zero-crossing frequencies, i.e., the frequencies at which the CTF reverses its phase, is key to recovering true phases and restoring higher resolution information. Traditionally, the Thon rings [3] observed in a power spectrum of a projection have been used for this purpose. However, the low dose in cryo-ET and the defocus gradient, particularly in highly tilted images, reduce Thon ring visibility and make the determination of defocus a challenge. If the specimen studied has an ordered structure, such as a 2-D crystal, the defocus can be accurately deduced from the diffraction patterns [4], but this approach cannot be applied to amorphous structures, especially if they are embedded in ice without a carbon background. Periodogram averaging [5], which involves computing and averaging the power spectra from many small subregions of an image, has been developed for these cases and is being extended in this paper.

Once the CTF is known, the true power spectrum can, in theory, be recovered by undoing the CTF modulation, except at frequencies where the CTF is zero. In 3-D cryo-ET, the spatially variant nature of the CTF and the low SNR complicate the correction for CTF. Amplitude and phase-only correction methods have therefore been developed [6–8]. Here, we have implemented a phase-only correction, flipping the phase of the FFT between appropriate CTF zeros. As a partial correction for the attenuation of information at higher resolution, the higher frequencies in the averaged subvolumes are also boosted by the inverse of the CCD camera's modulation transfer function (MTF).

We have applied our methods to tilt series from several specimens, including bovine papilloma virus (BPV) and microtubules decorated with Eg5 motor molecules [9–11]. An average virus particle, generated by extracting and aligning particles from a tomogram, was compared with a map at 2 nm resolution based on icosahedral symmetry [12] to demonstrate the efficacy of the CTF correction. The programs that we have developed have been incorporated into the tilt series processing sequence of the IMOD software package [13] (<http://bio3d.colorado.edu/imod>, version 4.0.12 or higher for all the features described here).

2. CTF determination

2.1. Background on CTF determination

In cryo-electron microscopy (cryoEM), researchers have relied on periodogram averaging, rotational averaging, and background subtraction to bring out oscillations in the measured power spectrum that can be used to determine image defocus [5,6,14]. Periodogram averaging is a power spectrum estimator developed in the field of multidimensional signal processing. In cryoEM, image regions with similar defocus are commonly subdivided into tiles whose power spectra are averaged to yield an estimate. The averaging reduces the noise but the division into tiles decreases the resolution of the estimate, i.e., the number of frequency points between zero and Nyquist, the limiting frequency in a digital image. The choice of the number and size of the tiles thus involves a tradeoff between smoothness and resolution. It is important that the tiles contain enough data points (thus sufficient resolution in the final estimate) that CTF oscillations can be resolved.

Fernandez et al. [6] used the Welch version of this estimator, in which tiles are half overlapped to increase tile number. To further increase the number of tiles, tiles were included from the entire tilt series, provided that their defocus differed from that of the untilted image by less than a chosen threshold. This approach was called strip-based periodogram averaging, and good results were reported. We have also found that this approach produced power spectra with greatly reduced noise, so we adopted it in our implementation.

Assuming negligible astigmatism and thus perfectly circular Thon rings, the 2D power spectrum produced by periodogram averaging may be rotationally averaged to create a 1D spectrum.

Background subtraction is another essential step in making the oscillations stand out, because the background is generally several orders of magnitude stronger than the oscillations. In the past, the background has usually been modeled by fitting a parametric curve to the local minima of the 1D power spectrum. For example, Fernandez et al. [6] modeled it by fitting a cubic spline to the set of points in the 1D power spectrum at the positions of the zeros of the theoretical CTF. Mallick et al. [15] used an empirical parametric form for the noise and envelope functions and solved for the parameters by a constrained optimization search. However, Mindell and Grigorieff [14] calculated the background as a low-pass filtered version of the power spectrum. Below, we describe a novel approach that is also effective.

2.2. Challenges

Several freely available CTF determination programs, including ACE [15], CTFFIND3 [14], and TOMOCTFFIND [6], have been tested on our tilt series. None of these programs produced consistent defocus estimations. Even TOMOCTFFIND, which is most suitable for dealing with the low SNR in cryo-ET, was not consistently successful. After examining the scaled background-subtracted power spectrum of TOMOCTFFIND, we found that the subtraction resulted in dramatically different oscillations in the power spectrum, depending on the value of assumed defocus. This indicates that the background estimated by curve fitting is not consistent and appropriate in this case. In general, these programs base their analysis on the location of more than one minimum in the power spectrum, whereas tilt series images taken at a defocus of 6 μm or less have very little signal and thus no detectable minima past the first zero of the CTF.

2.3. Noise floor subtraction

For cryo-ET in our laboratory, the electron dose incident to the specimen is $1\text{--}2\text{e}/\text{\AA}^2$ per image and the pixel size is 7 – 10 $\text{\AA}$; typically 25% to 50% of incident electrons are recorded in each image. Our images thus contain 12 to 100 electrons per pixel. The low electron count suggests that the majority of the noise is quantum noise and that its power spectrum can be used to approximate the background in CTF determination. Our experiments confirmed that this was the case.

The quantum noise images were collected without any specimen in the beam and with doses ranging from 0.1 to $10\text{e}/\text{\AA}^2$ at the level of the specimen. Their 1D power spectra were obtained using periodogram averaging and rotational averaging. The background power spectrum for a projection was interpolated from two of these 1D power spectra that were taken with doses closest to the mean dose recorded at the camera for a projection image. Figure 1 shows 2D and 1D power spectra for images at 0° tilt from two tilt series: the high dose one taken with a Tecnai F30 (FEI, Eindhoven, The Netherlands), the low dose one with a Tecnai F20. The bottom panels show the results of dividing by the noise power spectrum; we will refer to this operation as background subtraction because the logarithm of a divided spectrum is the same as the difference between the logarithm of the original and the noise spectrum. These curves approach zero, or at least become nearly flat, at high frequency (Figs. 1E, F). This shows that the noise power spectrum is an adequate approximation of the background of the CTF. Zhang et al. [16] have also used images without a specimen to estimate the background noise in images from carbon film when evaluating CCD camera performance. In Fig. 1, the oscillations from the CTF are visible in both background-subtracted spectra; they are particularly prominent on the left because of both the higher dose and the signal from the carbon film. On the right, they are still more prominent than in our typical tilt series because of the relatively high defocus.

2.4. Shifting and then averaging 1D power spectra

The power spectrum from a single image of a tilt series is typically too noisy to allow an accurate estimate of the location of even the first zero (Fig. 1F). To improve the SNR, spectra can be averaged over tiles from multiple images of a tilt series [6]. Figure 2A, B show such averages for the central tiles from images in the 30–45° range of our two example tilt series. The strip-based periodogram averaging does improve the SNR of the curves.

In some cases, the oscillations in a power spectrum obtained from strip-based periodogram averaging may still be hard to discern, due to noise. We have developed a way to accentuate the oscillations by averaging in more of the tiles from higher tilt images. The idea is that while the absolute defocus is unknown, if we have two power spectra calculated from two sets of tiles at different defocus, the known defocus difference between the two sets of tiles can be used to estimate the shift needed to align the first or the second zeros of the two curves. In order to reinforce the data at both zeros, the power spectrum of an off-center tile is shifted by a variable amount. Two shifts are computed, the ones that would align the first or the second zeros. Frequencies to the left of the first zero are shifted by the first amount, frequencies to the right of the second zero are shifted by the second amount, and frequencies between the two zeros are shifted by a linear combination of the two amounts. With these shifts, all the tiles in a tilt series can be included in computing the power spectrum. Figure 2C, D show the power spectra obtained after all the extra tiles in the 30° to 45° tilt range have been added to the curves shown in the top row of Fig. 2. As the figure shows, the curves become smoother and thus are easier to fit. An initial error in the defocus can limit the accuracy of the shifting and spread out the minimum in the curve, but if the minimum can be detected at all, the process can be iterated with a new estimated defocus, and the shifting will be more accurate on successive iterations. In practice, we have found that the defocus estimate converges in a few iterations. The changes in the curve upon iteration are small for high defocus but substantial for lower defocus (3–4 μm). In the latter case, it is helpful if the initial defocus estimate is accurate to within 0.5 μm , which can be achieved by starting with an estimate from only the central tiles or from a nearby range of tilt angles.

Being able to include all of the tiles at any tilt allows any subrange of tilts to be used, with the SNR of the power spectrum depending only on the number of images included and not on the tilt angle. It is thus not necessary to assume that the entire tilt series is at the same defocus. One can analyze a set of subranges to see if there are long-term trends in the defocus, such as might happen if the area used for focusing during the tilt series were at a different height than the area being recorded. Examples of such analyses appear below (Fig. 4, Fig 7, and Fig 9).

Power spectra must be computed from original rather than aligned tilt series images, because the interpolation needed to align an image distorts the power spectrum by attenuating the high frequencies. Thus, tiles are sorted into strips parallel to the tilt axis, which can be at any angle, and the spectrum shifts for a strip are based on its distance from the tilt axis.

2.5. User interaction and CTF fitting

The signal from the CTF was so weak for some of our tilt series that we saw the need for a user interface to allow visualization, verification, and control of the fitting process that finds the location of the first zero. Figure 3 shows the graphic user interface (GUI) of our CTF determination program. The vertical axis is the log of signal amplitude, and the horizontal is the normalized frequency, with 1.0 being the Nyquist frequency. The red curve is the 1D power spectrum obtained by noise floor subtraction. The green curve is fit to a portion of the red curve in order to determine the defocus. The number of projections included can be controlled by specifying a tilt angle range. The detected defocus for any given range of tilt angles can be stored in a text file and then be used by CTFPHASEFLIP.

The program provides three methods for fitting to a power spectrum, all with the goal of locating the first zero accurately. The most useful one fits to a CTF-like function given by:

$$b + se^{-kq} \text{CTF}(q, \Delta z)^p$$

Where:

b is a baseline value

s is a scale factor

k is an exponential decay constant representing an envelope function

q = spatial frequency

Δz = defocus

$\text{CTF}(q, \Delta z)$ is a CTF function similar to that given in Eq. 1 below, and

p = a power applied to the CTF function.

In fitting to this function, the goal is to get the best smooth curve through the dip at the first zero, not to fit a CTF curve over an extended frequency range, and the frequency limits of the fit should be adjusted accordingly. In the example of Fig. 3, the left limit has been set to the point beyond which the fitted curve diverges from the power spectrum. The right limit has been set about halfway between the first and second zeros, because the power spectrum falls off at this point, a pattern that we have seen in other low defocus data sets. Another reason to set the right limit before the second zero is to minimize the effects of a residual slope in the background-subtracted power spectrum. The power parameter, p , is limited to the range 0.5–2 and allows the curve to adapt to the width of the dip in the power spectrum. However, the parameter is optional because the extra degree of freedom can destabilize the fit if the dip is particularly weak or the power spectrum is especially noisy.

The second method fits a single polynomial to data around the first zero and locates the zero at the minimum of the polynomial; it can be used if the CTF-like fitting fails, and if there is a discernable dip at the first zero. The third method involves fitting two separate curves (either a straight line or a Gaussian) before and after the first zero and taking the intersection of the two curves as the first zero. This method is less robust because it is based on the intersection between curves at or beyond the frequency range to which they were fit, where they are least accurate; however, it can be useful if there is no discernable dip at the first zero.

Once the parameters for CTF-like curve-fitting have been set on a representative subrange of the tilt series, they can be applied to obtain consistent defocus estimates from other ranges of angles. Figure 4 shows 6 examples in which the CTF fitting was applied in this way for an angular range of 20° (or 10° in one case) at intervals of 10°, generally without changing the fitting parameters. All tilt series were acquired with the SerialEM program developed in our laboratory [17]. Overall defocus variation ranged from 0.25 to 1 μm . The low jitter in the curves from point to point shows that the curve fitting is fairly robust, even for these relatively small angular ranges. The total dose for these data sets was 80–110 $\text{e}^-/\text{\AA}^2$, and thus the dose contributing to each power spectrum being analyzed was typically $\sim 15 \text{ e}^-/\text{\AA}^2$. Tilt series taken at a lower total dose would require a correspondingly larger angular range to achieve the same SNR in the power spectra but it should still be possible to analyze the tilt series in three subranges.

3. CTF correction

Two distinct methods have been developed for the correction of tilted images with a spatially variant CTF: the localized line-by-line approach of Fernandez et al. [6] and the global approach of Winkler and Taylor [7] and Philippsen et al. [8]. The method of Winkler and Taylor [7] involves taking a 1D Fourier transform along the axis parallel to the tilt axis and then solving a matrix equation of Fourier components at each frequency component along that axis. Philippsen et al. [8] showed that the spatially variant CTF for a weak-phase specimen is a linear transform in the frequency domain and proposed to invert the transform by solving a system of linear equations. Each global approach requires solving equations in frequency space, and in our tests with Winkler and Taylor's [7] method it was sensitive to initialization. In the approach that separately corrects each line parallel to the tilt axis, a strip of image centered on the line is corrected, as if it were all at the same defocus, and the line is then replaced by the central line of the strip. The accuracy of this approach depends on the amount of defocus variation within a strip, which determines how much the first zero varies within the strip and the degree to which phase is flipped inappropriately for some image components. However, there is a tradeoff between having the strip narrow enough to minimize defocus variation and wide enough to provide adequate resolution (number of pixels) in that direction in frequency space. In CTFPHASEFLIP, this tradeoff is resolved by requiring strips to have a certain minimum width, even if that makes defocus vary by more than a set amount. To assess the extent of inappropriate phase flipping engendered by this approach, we start with a standard equation for the phase contrast transfer for a weak phase object [18]:

$$B(\theta^*) = -2\sin\left[\frac{\pi}{2}(\theta^{*4} - 2\theta^{*2}\Delta z^*)\right] \quad (1)$$

Where Δz^* is reduced defocus and θ^* is reduced spatial frequency, defined by:

$$\begin{aligned} \Delta z^* &= \Delta z / \sqrt{C_s \lambda} \\ \theta^* &= (C_s / \lambda)^{1/4} \lambda q \\ \Delta z &= \text{defocus} \\ C_s &= \text{spherical aberration} \\ \lambda &= \text{wavelength of electrons (0.00197 nm at 300 kV, 0.00251 nm at 200 kV)} \\ q &= \text{spatial frequency} \end{aligned}$$

The first zero occurs when the argument of the sine in Eq. 1 reaches π , and the equation is easily solved for the value of θ^* at the first zero. The resulting frequency is:

$$q_0 = \sqrt{\Delta z^* - \sqrt{\Delta z^{*2} - 2/(\lambda(C_s/\lambda)^{1/4})}} \quad (2)$$

The numerator of Eq. 2 is very close to $2/\sqrt{\Delta z^*}$ for high values of Δz^* corresponding to greater than 1 μm underfocus. This expression yields a simple approximation,

$$q_0 \approx \frac{2}{\sqrt{\lambda \Delta z}}$$

To determine the degree to which there is inappropriate phase-flipping of frequency components when correcting a strip of data, we used Eq. 2 to compute the amount that the first zero shifts from the middle to the edge of a strip in some extreme cases. The values in the next

to last column of Table 1 show that even at high tilt angles, the first zero shifts by less than one pixel in the FFT of the strip, with a shift usually around 0.5 pixel or less. Note that only the image components near the edge of the strip have frequency terms inappropriately flipped. Thus, it seems intuitively that the predominant effect of the inappropriate flipping would be on those components. If that is the case, then the effect on the central line being corrected should be minimal.

One drawback of the line-by-line approach is that it needs to compute the FFT of a strip to correct each line parallel to the tilt axis, which is time-consuming. We have therefore implemented a method to make the computation more efficient. The projection is divided into overlapping strips as shown in Figure 5, and only the FFTs of these strips are computed. By correcting strips at intervals of about 20 pixels, the computation is much faster than with the line-by-line method. To avoid discontinuities in transitioning from one strip to the next, the value for a pixel located between the central lines of two overlapping strips is obtained by interpolating between its values in each of the strips, with the interpolation fraction varying linearly from one central line to the next. As shown by the difference image in Figure 6, this strip-based method gives essentially the same result as the line-by-line approach. Moreover, the last column of Table 1 shows that the amount of variation in the position of the first zero between adjacent strips is minimal (typically less than 0.1 pixel in the FFT), even at high tilt.

Our correction method simply flips the phase in the FFT at frequencies between successive odd and even zeros of the CTF. Thus far, we have not implemented Wiener-like amplitude correction near the zeros, because in our tests, it amplified noise and further reduced the already low SNR in the original image.

4. Improvements from CTF correction

The effects of CTF correction are most evident when multiple copies of a structure can be averaged to increase image SNR at the high frequencies being corrected. We have examined several structures before and after CTF correction, using the PEET software developed in our laboratory [19] for aligning and averaging particles extracted from tomograms. To average bovine papilloma virus (BPV), we used a tomogram from a tilt series taken at 9 μm underfocus, marked 519 virus particles, and aligned them to one particle chosen as a reference, searching with an angular range of 36° and increment of 3° around each axis. A new reference was generated by averaging the 350 particles with the highest correlations, and the alignment was refined with an angular increment of 1° . We found angles for rotating each of the 12 five-fold vertices of the resulting average to the z-axis and combined these angles with the alignment angles to provide a starting point for a search that included each particle in 12 different orientations. Six iterations of refinement were done, ending with an angular increment of 0.6° ; each iteration generated a new reference from 3000 particles sampled equally from the 12 sets of orientations. The resulting alignment was also used to generate an average from the tomogram, based on the CTF-corrected tilt series. We also tried refining the alignment of CTF-corrected particles to the corrected average, but there was no improvement in the Fourier shell correlation (FSC), a standard method for assessing resolution of averages [20,21]. Figs. 7A and B show uncorrected and corrected averages of the 4000 particles with the highest correlation coefficients, regardless of which orientation sets they belonged to.

It would be convenient if the benefit from CTF correction could be revealed by an increase in the standard FSC, but it cannot. Apparently, the phase-inverted information past the first zero correlates well between the averages from two halves of an uncorrected data set, probably because it is equivalently incorrect in both halves. It is thus necessary to compare an average structure from tomograms with a higher resolution map obtained by other means.

We have demonstrated the efficacy of CTF correction by computing the FSC between the BPV averages and a map at 2 nm resolution based on a reconstruction from single particle cryo-EM [12] (Fig 7F). Figure 8 shows clearly that information between the first and second zero of the CTF had an inverted phase in the average from uncorrected data, and that the CTF correction appropriately flipped the phase there. Note, however, that in order to see these effects, it is necessary to compute the FSC, using the real component of the conjugate product of corresponding Fourier components, as in Frank (2006), rather than using the magnitude of the product, as in Harauz and van Heel (1986). Even the weak data between the third and fourth zeros (at 0.35 and 0.41 nm⁻¹, respectively) has had its phase restored. Although the FSC in this range is much too low for image details to be significant at these frequencies, there is apparently sufficient statistical power to extract one bit of information from all of the data in a spherical shell, namely whether the phase has been restored by the correction.

We have also used these comparisons between our BPV average and the 2-nm map to assess how sensitive the CTF correction is to errors in the defocus estimate. Using a series of different defocus values, we repeated the CTF correction of the tilt series and recomputed the tomogram, the average of aligned particles, and the FSC with the 2-nm map. The results are shown in Figs. 8B and C for the four frequency components that varied the most: ones near the first, second, and third zeros. Three of the four curves have their peaks at the measured defocus, and the other is only slightly higher 0.25 μm away, indicating that the defocus values picked from the power spectra (Fig. 7E) were appropriate for correcting the data. The curves fall off appreciably farther than 0.25 μm away from their peaks, indicating that defocus estimates should be accurate to within 0.25 μm for the most effective correction, at least in this range of underfocus. Finally, note that even the weak correlation past the third zero varies consistently with the assumed defocus, indicating again that the correction has been appropriate out to the fourth zero.

Two other experiments were done with these averages. First, we computed the CTF correction using a single defocus for the whole tilt series, rather than values determined from three tilt ranges. The FSC was consistently worse from the first to the fourth zero, but only by small amounts (up to 0.008). Thus, correcting for a varying defocus is appropriate but may have only minor benefit. Second, we assessed whether correcting with a finer-grained variation in defocus would be helpful. We used the CTF-like fitting to estimate defocus on successive, non-overlapping pairs of images and corrected the tilt series with these defocus values. The resulting FSC was worse beyond 0.25 nm⁻¹ by up to 0.014. Thus, the inaccuracy of these defocus estimates from noisy power spectra outweighs any benefit in correcting defocus variations on a finer scale.

The improvement produced by CTF correction alone is quite subtle in the averages of Figs. 7A and B because high frequencies are strongly attenuated by a variety of factors. The FSC is insensitive to this attenuation because it measures whether information is correlated at a particular frequency, but not whether it has the proper relative amplitudes across a range of frequencies. For example, the ratio of the magnitudes in our average of BPV and in the 2-nm map (Fig. 7F) falls 20-fold from its value at low frequency to its value at 5 nm resolution, where the FSC is still 0.7.

To provide a partial compensation for the attenuation of frequencies, we have inverse-filtered the averages with the measured modulation transfer function (MTF) of the CCD camera. The MTF was computed from an image of a sharp edge (the beam stop), using a program that finds the average intensity profile perpendicular to the edge and takes the Fourier transform of the derivative of this profile. The program MTFILTER was used to filter by the inverse of this MTF; this program provides for a cutoff frequency at which the inverse is attenuated back to 1 by multiplying by a Gaussian with a specified sigma. In the case of the BPV averages, the

cutoff corresponded to a resolution of 2.8 nm and the inverse rose to 6 before being attenuated. Inverse filtering makes the effect of CTF correction more obvious for the BPV averages in Figs. 7C and D (see legend for details). It also enhances important details in the averages from a microtubule fully decorated with Eg5 motor domains, presented next.

To average the microtubule, we first marked points along the center of the microtubule and along several protofilaments. These trajectories were used to pick particle center points every ~8 nm and to compute initial rotations to follow the twist of the microtubule. The alignment search on 244 particles initially used one particle as a reference and allowed rotation up to 18° about the microtubule axis, then used three iterations with constrained searches about all three axes to refine the alignment. Because the MT had 15 protofilaments and lacked a seam, we then used a search that included each particle 15 times, rotated by 24° increments, refining the alignment in 3 iterations ending with an angular increment of 0.6°. The new reference each time included 93% of the particles in order to prevent any bias due to orientation with respect to the tomographic missing wedge. The alignment was done on a CTF-corrected tomogram and applied to particles from an uncorrected tomogram.

The averages from the Eg-5 decorated MT appear in Fig. 9. Again, the differences are much clearer with the high frequencies boosted by MTF correction. The Eg5 heads are shaped differently after CTF correction; and the tubulin subunits are displaced by phase inversion in the uncorrected average.

5. Discussion

This paper has described methods implemented in IMOD for determining defocus and correcting the effects of microscope CTF and camera MTF in tomographic tilt series of vitrified specimens. The determination of defocus is more difficult for cryo-ET than for single-particle analysis, primarily because the electron dose for any one image in cryo-ET is typically 10–50 fold lower than for single-particle analysis. The task is even harder when the defocus and image pixel size are chosen to place the first zero of the CTF near half-Nyquist (as is often done in our laboratory), for then there is very little information about higher order zeros. The paucity of signal from higher order zeros motivated us to implement subtraction of an empirical noise background instead of trying to find a background by fitting to multiple local minima. This subtraction makes the power spectrum close enough to flat at high frequencies that it becomes much easier to see the first zero and analyze the curve for its location. This method of flattening the spectrum has an advantage over filtering of the power spectrum in that it preserves the structurally derived signal, whereas filtering inevitably removes some of this signal along with the noise background. We experimented with such filtering of the 1D spectrum (either high-pass filtering, subtraction of a low-pass filtered spectrum, or subtraction of a box convolution average [14]) and found that the filtering distorted the curve enough to shift the location of the first zero, leading to errors of 0.5–1 μm in the defocus estimate.

Because our experience shows that it is not feasible to analyze the defocus for individual images, our program combines data from multiple images to estimate the power spectrum, as was done by Fernandez et al. [6]. We developed a method for combining data from the different z heights in an image. As a result, all of the available image data can be used, regardless of the tilt angle. This capability allows the analysis of defocus in subsets of images, even at high tilt. The ability to detect systematic trends in defocus through the tilt series is potentially important, because the area used for focusing in low dose image acquisition could be at a different height, and this difference would change systematically as the specimen tilts. As shown above (Fig. 4), some tilt series do show such trends, with the defocus varying by up to 1 μm . The trends in defocus are increasingly significant at lower defocus and it is thus useful that they can be detected and their effects accounted for in our correction procedure.

Defocus is determined here primarily from the position of the first zero, because this is usually the only detectable zero in power spectra from our tilt series. This method limits the reliability of the CTF correction between higher order zeros. Our experiments with the BPV data set indicate that when the defocus is accurate to within 0.25 μm , corrections are reliable out to the fourth zero. For noisier data where the accuracy of determination is less than this, the CTF correction may be effective only for flipping the phase between the first and second zero.

Zanetti et al. [22] recently explored the effect of errors in defocus estimation on the resolution of averaged particles from CTF-corrected tomograms and found that errors up to 0.5 μm had little effect on resolution, consistent with the results shown above. They also found that correcting with relatively inaccurate defocus estimates for individual images was no better than using a single defocus value for the whole tilt series, consistent with our finding that the inaccuracy of estimates from single images slightly degrades the resolution of averages.

Our implementation of CTF correction takes a series of narrow strips from the image, corrects each one for a uniform CTF, and uses only a subset from the center of each strip. This method should be sufficiently accurate for the range of resolutions currently attainable with cryo-ET (2–5 nm). As described above, the method flips phases inappropriately only for information from the edges of strips at high tilt. In support of this point, Philippsen et al. [8] calculated that the error from correcting a 100-nm wide strip with a uniform CTF is minimal for resolutions of better than 2 nm.

We found that CTF correction produced very subtle changes in averaged structures from tomograms unless the attenuation of information at high frequencies was compensated to some extent. In conventional single-particle reconstructions, a variety of factors contribute to the loss of amplitudes at high frequencies; their effects are typically modeled with attenuation by $\exp(-Bq^2)$, where q is frequency and B is the “temperature factor” [23,24]. The B -factor is obtained by fitting to the decay of power spectra at high frequencies, and its predominant effect is at those high frequencies; the highest B -factors found, $\sim 500 \text{ \AA}^2$ [23], account for an attenuation of only 0.78 at $1/3 \text{ nm}^{-1}$. This approach is thus not helpful for lower resolution data from cryo-ET. Rather than trying to account for all the factors that attenuate amplitudes in our data, we chose instead to correct for the known effects of the MTF of the CCD camera. These effects are substantial over the middle range of frequencies; e.g., with a pixel size of 0.7 nm, the MTF for the camera on our F20 falls to 0.8 at $1/10 \text{ nm}^{-1}$ and to 0.25 at $1/3 \text{ nm}^{-1}$. Multiplying by the inverse of the MTF can thus substantially improve the degree to which an averaged volume represents the actual structure, and we recommend using this correction in the absence of a more sophisticated approach. We choose a frequency at which to attenuate the multiplication so as to avoid excessive amplification of noise in the corrected volume; it would be possible instead to attenuate it with a weighting factor based on the FSC curve [25].

Before we began correcting for the MTF, we saw little effect from CTF correction and were dubious about its value. The examples shown here were taken at 8–9 μm , a significantly higher defocus than we would normally use, expressly to emphasize the effects of the CTF. With MTF correction, however, we have seen appreciable effects from CTF correction of tilt series taken with less defocus. Indeed, whenever the resolution of an average structure extends beyond the first zero of the CTF, it is important to use CTF correction to recover the higher frequency information.

Acknowledgments

We thank J. Richard McIntosh for reading the manuscript. This work was supported by Grant Number P41RR00592 to A. H. Hoenger from the National Center for Research Resources (NCRR) and by R01 EB005027 to D. N. Mastrorade from the National Institute of Biomedical Imaging and Bioengineering (NIBIB), both components of the

National Institutes of Health (NIH). Its contents are solely the responsibility of the authors and do not necessarily represent the official view of NCRR or NIH.

References

1. Zhu J, Penczek PA, Schroder R, Frank J. Three-dimensional reconstruction with contrast transfer function correction from energy-filtered cryoelectron micrographs: procedure and application to the 70S Escherichia coli ribosome. *J. Struct. Biol* 1997;118:197–219. [PubMed: 9169230]
2. Toyoshima C, Unwin N. Contrast transfer for frozen-hydrated specimens: determination from pairs of defocused images. *Ultramicroscopy* 1988;25:279–291. [PubMed: 3188279]
3. Thon F. Zur Defokussierungsabhängigkeit des Phasenkontrastes bei der elektronenmikroskopischen Abbildung. *Z. Naturforschung* 1966;476–478.
4. Henderson R, Baldwin JM, Downing KH, Lepault J, Zemlin F. Structure of purple membrane from halobacterium-halobium, recording, measurement and evaluation of electron-micrographs at 3.5 Å resolution. *Ultramicroscopy* 1986;19:147–178.
5. Fernandez J-J, Sanjurjo JR, Carazo J-M. A spectral estimation approach to contrast transfer function detection in electron microscopy. *Ultramicroscopy* 1997;68:267–295.
6. Fernandez JJ, Li S, Crowther RA. CTF determination and correction in electron cryotomography. *Ultramicroscopy* 2006;106:587–596. [PubMed: 16616422]
7. Winkler H, Taylor KA. Focus gradient correction applied to tilt series image data used in electron tomography. *J. Struct. Biol* 2003;143:24–32. [PubMed: 12892723]
8. Philippsen A, Engel HA, Engel A. The contrast-imaging function for tilted specimens. *Ultramicroscopy* 2007;107:202–212. [PubMed: 16989948]
9. Krzysiak TC, Wendt T, Sproul LR, Tittmann P, Gross H, Gilbert SP, Hoenger A. A structural model for monastrol inhibition of dimeric kinesin Eg5. *Embo J* 2006;25:2263–2273. [PubMed: 16642039]
10. Sawin KE, LeGuellec K, Philippe M, Mitchison TJ. Mitotic spindle organization by a plus-end-directed microtubule motor. *Nature* 1992;359:540–543. [PubMed: 1406972]
11. Turner J, Anderson R, Guo J, Beraud C, Fletterick R, Sakowicz R. Crystal structure of the mitotic spindle kinesin Eg5 reveals a novel conformation of the neck-linker. *J Biol Chem* 2001;276:25496–25502. [PubMed: 11328809]
12. Trus BL, Roden RB, Greenstone HL, Vrhel M, Schiller JT, Booy FP. Novel structural features of bovine papillomavirus capsid revealed by a three-dimensional reconstruction to 9 Å resolution. *Nat. Struct. Biol* 1997;4:413–420. [PubMed: 9145113]
13. Kremer JR, Mastronarde DN, McIntosh JR. Computer visualization of three-dimensional image data using IMOD. *J. Struct. Biol* 1996;116:71–76. [PubMed: 8742726]
14. Mindell JA, Grigorieff N. Accurate determination of local defocus and specimen tilt in electron microscopy. *J. Struct. Biol* 2003;142:334–347. [PubMed: 12781660]
15. Mallick SP, Carragher B, Potter CS, Kriegman DJ. ACE: automated CTF estimation. *Ultramicroscopy* 2005;104:8–29. [PubMed: 15935913]
16. Zhang P, Borgnia MJ, Mooney P, Shi D, Pan M, O'Herron P, Mao A, Brogan D, Milne JL, Subramaniam S. Automated image acquisition and processing using a new generation of 4K × 4K CCD cameras for cryo electron microscopic studies of macromolecular assemblies. *J Struct Biol* 2003;143:135–144. [PubMed: 12972350]
17. Mastronarde DN. Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol* 2005;152:36–51. [PubMed: 16182563]
18. Reimer, L. *Transmission Electron Microscopy : Physics of Image Formation and Microanalysis*. Vol. 4th ed.. Berlin: Springer; 1997.
19. Nicastro D, Schwartz C, Pierson J, Gaudette R, Porter ME, McIntosh JR. The molecular architecture of axonemes revealed by cryoelectron tomography. *Science* 2006;313:944–948. [PubMed: 16917055]
20. Harauz G, van Heel M. Exact filters for general geometry three dimensional reconstruction. *Optik* 1986;73:146–156.
21. Saxton WO, Baumeister W. The correlation averaging of a regularly arranged bacterial cell envelope protein. *J. Microsc* 1982;127:127–138. [PubMed: 7120365]

22. Zanetti G, Riches JD, Fuller SD, Briggs JAG. Contrast transfer function correction applied to cryo-electron tomography and sub-tomogram averaging. *J. Struct. Biol.* 2009In Press
23. Frank, J. Three-dimensional Electron Microscopy of Macromolecular Assemblies. Vol. 2nd ed.. Oxford: Oxford University Press; 2006.
24. Glaeser RM, Downing KH. Assessment of resolution in biological electron crystallography. *Ultramicroscopy* 1992;47:256–265. [PubMed: 1481279]
25. Rosenthal PB, Henderson R. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J. Mol. Biol* 2003;333:721–745. [PubMed: 14568533]

Figure 1.

Power spectra for images from tilt series taken at a high dose over a carbon support film (left column: dose 10 e⁻/Å², 10 μm underfocus, ice-embedded specimen) and a low dose over a hole in the film (right column: dose 1 e⁻/Å², 9 μm underfocus). A, B: FFTs of 0° tilt images. C, D: Logarithms of the original and background power spectra for tiles from these images. E, F: Logarithms of the original power spectrum divided by the background power spectrum.

Figure 2.

Rotationally averaged power spectra obtained using tiles from images at tilt angles from 30° to 45° in the high and low dose tilt series used for Figure 1. A, B. Power spectra based on tiles whose defocus is within 200 nm of the defocus at the center of the image. C, D. Power spectra resulting from adding in tiles at higher defocus difference, after shifting their power spectra as described in the text.

Figure 3.

User interface of the CTFPLOTTER program. The red curve is the averaged power spectrum, whose computation is controlled by the parameters in the right dialog box. The only text fields that the general user needs to modify are the expected defocus and the limiting tilt angles; the rest are available for experimentation. The radio buttons control whether to compute the power spectrum from just central tiles or from all tiles, and which defocus to assume when shifting off-center power spectra into register with central ones. Using the current defocus estimate allows one to iterate rapidly until the estimate converges. The green curve is fit to the power spectrum, based on the parameters in the left dialog box, which allows one to select the kind of curve(s) to fit and the range of frequencies over which to fit (see text for details). The “Z” and “D” outputs at the top of the window show the relative frequency and defocus of the first zero. These values are set each time that curve-fitting is done, but can be changed by clicking at a desired x coordinate. If the position of the second zero is clear, the user can click on it with a different mouse button and the “D2” and “D-avg” outputs will show the defocus based on that zero and the average of that with the defocus based on the first zero.

Figure 4.

Plots of defocus versus tilt angle for 6 tilt series, showing that systematic changes in defocus do occur and can be detected with our approach. Values were obtained by fitting a CTF-like curve to power spectra from a 20° or 10° angular range at 10° intervals. The frequency fitting range was set from an initial power spectrum over a 40° range around zero tilt and then applied to all other angular ranges, requiring modification in only one case. The tilt series were of BPV and taken on the F20 (filled squares, open circles); of Eg5-decorated microtubules taken on the F20 (filled triangles); and of the ventral disk of *Giardia* taken on the F30 with energy filtering (filled circles, open triangles, crosses). One of the latter had such a good CTF signal that it was analyzed successfully with 10° ranges (filled circles). The entire angular range was

analyzed for all sets except the low-defocus BPV set (open circles), where the power spectrum was too slanted in the last subrange.

Figure 5.

Representation of adjacent strips used to speed up the line-by-line approach for correcting CTF. Strips 1 and 2 are transformed, corrected, and inverse transformed; these data are used to correct all lines between the center lines of the two strips. The lines are parallel to the tilt axis, which is vertical in the aligned images that are corrected.

Figure 6.

Difference between correcting using a line-by-line method and using strip-based correction with interpolation between strips. A) Original projection image of 10-nm colloidal gold particles on a carbon film tilted to 65°. The image was taken on an F30 at 5.7 μm underfocus and the region shown is at 6.5 μm defocus. The tilt axis is horizontal in these images. B) Image corrected one line at a time. The effect of the correction is difficult to see when comparing images side-by-side, but the underfocus fringes around the gold particles are noticeably reduced (arrows). (The fringes remaining after correction reflect the dominance of frequencies below the first zero of the CTF.) C) Image corrected by interpolating between corrected strips 20 pixels apart. D) Difference between the images in B and C, showing no organized features

related to the structure. The difference is less than 0.5% of the original intensity range, and it would appear uniformly gray if shown at the same intensity scaling as the images.

Figure 7.

Averages of bovine papilloma virus (BPV) from a tilt series taken at 9.6 μm underfocus, based on 4000 aligned orientations of a subset of 519 virus particles. Each panel shows an image slice near the equator of the virus and an inset with a surface rendering of a capsomere at a 5-fold vertex. A. Average from tomogram without CTF correction. B. Average from tomogram after CTF correction of the tilt series. C. Average without CTF correction, filtered by the inverse of the camera MTF to restore high frequency information. D. Average with CTF correction and inverse filtering by the MTF. F. Map at 2-nm resolution [12] used for Fourier shell correlation with the averages. Black arrows mark capsomeres that are more appropriately shaped after CTF correction; the difference is slight between A and B but more noticeable

between C and D. White arrows point to the vertices of the pentamer at the 5-fold vertex; the orientation of the pentamer is clearly incorrect in C. E. Power spectra used to determine the defocus for CTF correction for three angular subranges of the tilt series, and the power spectrum based on the whole tilt series. The curves are displaced vertically by arbitrary amounts. Vertical lines mark the frequencies manually chosen for the zeros, corresponding to 9.55, 9.6, and 9.85 μm underfocus for -56° to -16° , -20° to 30° , and 25° to 65° , respectively. The scale bar in F corresponds to 10 nm in the main images and 6 nm in the insets.

Figure 8.

A. Fourier shell correlation between averages of BPV from tomograms and the 2-nm map, showing the phase inversion before CTF correction. The FSC is computed using the real component of the conjugate product of corresponding Fourier components. Vertical dashed lines show the locations of the first four zeros of the CTF. B, C. FSC values at 4 spatial frequencies near CTF zeroes, as a function of the defocus assumed when computing the CTF correction. The curves indicate that the best correction overall was obtained with the measured defocus, and that an error of 0.25 μm has relatively little impact on the value of the correction.

Figure 9.

Averages from a microtubule decorated with Eg5, obtained by aligning and averaging subunits from a microtubule in a tomogram generated from a tilt series taken at 7.4 μm underfocus. A. Average without CTF correction. B. Average with CTF correction of tilt series. C. Uncorrected average filtered by the inverse of the CCD camera MTF. D. Inverse-filtered, corrected average. In addition to the obvious differences in the appearance of the Eg5 heads, note that both tubulin subunits show up clearly after both corrections in D (arrows) but are less distinct before CTF correction and appear displaced in A and C because of the phase inversion of 4-nm information. E. Power spectra used to determine the defocus for CTF correction, as in Fig 7. Vertical lines mark the frequencies manually chosen for the zeros, corresponding to 7.15, 7.4, and 7.5 μm underfocus for -65° to -25° , -28° to 12° , and 10° to 52° , respectively. F. Fourier shell correlation between averages based on two halves of the data; vertical lines show the location of CTF zeros. Bar = 10 nm.

Table 1

Variation in first zero of CTF in tilted strips

Defocus (μm)	Tilt angle ($^\circ$)	Pixel size (nm)	Strip width (pixels)	Δz : middle to edge (nm)	Δq_0 : middle to edge (pixels)	Δq_0 : between strips (pixels)
6	60 $^\circ$	1.0	128	128.0	0.39	0.06
6	60 $^\circ$	0.7	143	100.0	0.24	0.03
6	70 $^\circ$	1.0	128	187.1	0.57	0.09
6	70 $^\circ$	0.7	128	131.0	0.28	0.04
3	60 $^\circ$	0.7	143	100.0	0.67	0.09
3	60 $^\circ$	0.5	200	100.0	0.67	0.07
3	70 $^\circ$	0.7	128	131.0	0.78	0.12
3	70 $^\circ$	0.5	137	100.0	0.46	0.07

The amount of shift in the first zero of the CTF, Δq_0 , is computed for different conditions of defocus, tilt angle, and pixel size, assuming the default parameters for running CTFPHASEFLIP. The strip width is set so that the defocus change across a strip will be no more than 200 nm, but it is then constrained to be between 128 and 256 pixels. Δz is the change in z height from the middle to the edge of a strip, and the next column, Δq_0 is the corresponding change in location of the first zero, expressed as pixels in the FFT. The final column shows the shift in q_0 between adjacent strips separated by 20 pixels. The computation assumes 300 kV electrons; shifts are about 15% less at 200 kV.