Published in final edited form as:

J Struct Biol. 2016 August ; 195(2): 207–215. doi:10.1016/j.jsb.2016.06.003.

An Algorithm for Estimation and Correction of Anisotropic Magnification Distortion of Cryo-EM Images without Need of Pre-calibration

Guimei Yu¹, Kunpeng Li¹, Yue Liu¹, Zhenguo Chen¹, Zhiqing Wang¹, Rui Yan¹, Thomas Klose¹, Liang Tang², and Wen Jiang^{1,#}

¹Markey Center for Structural Biology, Department of Biological Science, Purdue University, West Lafayette, IN, USA

²Department of Molecular Biosciences, University of Kansas, Lawrence, KS, USA

Abstract

Anisotropic magnification distortion of TEM images (mainly the elliptic distortion) has been recently found as a potential resolution-limiting factor in single particle 3-D reconstruction. Elliptic distortions of ~1–3% have been reported for multiple microscopes under low magnification settings (e.g., 18,000x), which significantly limited the achievable resolution of single particle 3-D reconstruction, especially for large particles. Here we report a generic algorithm that formulates the distortion correction problem as a generalized 2-D alignment task and estimates the distortion parameters directly from the particle images. Unlike the present pre-calibration methods, our computational method is applicable to all datasets collected at a broad range of magnifications using any microscope without need of additional experimental measurements. Moreover, the per-micrograph and/or per-particle level elliptic distortion estimation in our method could resolve potential distortion variations within a cryo-EM dataset, and further improve the 3-D reconstructions relative to constant-value correction by the pre-calibration methods. With successful applications to multiple datasets and cross-validation with the pre-calibration method, we have demonstrated the validity and robustness of our algorithm in estimating the distortion; Correction of the elliptic distortion significantly improved the achievable resolutions by ~1–3 folds and enabled 3-D reconstructions of multiple viral structures at 2.4–2.6 Å resolutions. The resolution limits with elliptic distortion and the amounts of resolution improvements with distortion correction were found to strongly correlate with the product of the particle size and the amount of distortion, which can help assess if elliptic distortion is a major resolution limiting factor for single particle cryo-EM projects.

[#]Corresponding author: jiang12@purdue.edu.

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.

Keywords

anisotropic magnification distortion; elliptic distortion; *jspr*; generalized 2-D alignment; single particle 3-D reconstruction

Introduction

Development of direct electron detectors (DDD) with superior electron detective quantum efficiency allows fast speed recording of movies to correct beam-induced motions of particles. This has significantly boosted the achievable resolutions of single particle cryo-Electron Microscopy (cryo-EM) of a wide range of structures including small protein complexes, to near-atomic resolutions (Bai et al., 2015; Banerjee et al., 2016; Bartesaghi et al., 2015; Clemens et al., 2015; Jiang et al., 2015; Liao et al., 2013). Application of DDDs with small physical pixels sizes (e.g., 5 μm for Gatan K2 Summit) makes it possible to use lower nominal magnifications (e.g., 18,000x) of transmission electron microscopes (TEMs) than those (e.g., 50,000x or higher) previously used for photographic film or CCD for near-atomic resolution cryo-EM data collection (Guo et al., 2014; Liu et al., 2010; Liu et al., 2016; Yu et al., 2008; Zhang et al., 2010; Zhang et al., 2008). However, it has been observed recently that under these lower magnification settings of many microscopes the anisotropic magnification distortion caused by imperfect projector lenses could be a significant issue for high-resolution single particle cryo-EM work (Grant and Grigorieff, 2015c). With the presence of anisotropic magnification, the object will be magnified differently along different directions, thereby causing distortion of the projection images of target macromolecules (Capitani et al., 2006). The dominant distortion has been approximated as elliptic distortion (Capitani et al., 2006) by which a circle is transformed into an ellipse. The deteriorating effects of elliptic distortion are proportional to the size of the imaged objects. Therefore, it is a more severe problem for high-resolution cryo-EM of large macromolecules (e.g., viruses), which is probably one reason why viral samples have lost their leading role in high-resolution single particle cryo-EM in recent years (Banerjee et al., 2016; Bartesaghi et al., 2015; Campbell et al., 2015; Liao et al., 2013).

Current elliptic distortion estimation approaches all require additional experimental calibration for individual microscopes at each magnification using polycrystalline standards (Capitani et al., 2006; Grant and Grigorieff, 2015c; Hou, 2008; Mugnaioli et al., 2009; Zhao et al., 2015). The calibrations need to be remeasured after major services of the microscope or the camera, and even performed periodically to ensure the validity of the calibrations. In contrast, here we present a computational method to directly discover the elliptic distortion information from cryo-EM datasets. Our algorithm can computationally determine the amount and angle of the distortion without need of experimental calibrations. Moreover, as a data-adaptive method, our algorithm can be applied to any single particle cryo-EM dataset (old or newly collected), regardless of the facilities and settings used for data collection.

Methods

1. Quantification of anisotropic magnification distortion

The elliptic distortion of images can be represented as a general form of scaling transform of the image (Equation 1):

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} S_x & 0 \\ 0 & S_y \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = S_{avg} \begin{bmatrix} \frac{1}{1+\frac{\Delta}{2}} & 0 \\ 0 & 1+\frac{\Delta}{2} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad (1)$$

in which S_x and S_y represent the magnifications along X and Y directions, respectively. S_x and S_y can be described by an average magnification (S_{avg}) modulated by differential magnifications ($\Delta > 0$) and rewritten as $S_x = S_{avg}/(1 + \Delta/2)$, $S_y = S_{avg}(1 + \Delta/2)$. This general form of scaling is reduced to isotropic magnification when there is no distortion (i.e. $S_x = S_y$

$= S_{avg}$ when $\Delta = 0$) (Fig. 1A). Here we have adopted the formulation $\begin{bmatrix} \frac{1}{1+\frac{\Delta}{2}} & 0 \\ 0 & 1+\frac{\Delta}{2} \end{bmatrix}$ to ensure that the transform is area preserving ($S_x \cdot S_y \equiv S_{avg}^2$ for arbitrary Δ) and therefore does not change the mean magnification and sampling rate of images. In this definition, the Y-axis with a larger magnification (i.e. more elongated) is defined as the major axis of distortion. Equation 1 represents the simplified case when the major axis of distortion is along Y and the minor axis is along X-axis (Fig. 1B).

If the minor axis is at an arbitrary angle θ ($0^\circ < \theta < 180^\circ$ when the 2-fold symmetry of an ellipse is considered) (Figure 1C), Equation 1 needs to be extended to account for the angular offset of the distortion. Mathematically, elliptic distortion along an arbitrary angle θ is equivalent to a three-step process: rotate the image by $-\theta$ to align the minor axis along X axis; perform anisotropic scaling as shown in Equation 1; then rotate the image by θ to bring back the image to its original angular view. Since the 2-D in-plane rotation matrix is

$$R = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, \text{ Equation 1 is thus extended to}$$

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = S_{avg} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \frac{1}{1+\frac{\Delta}{2}} & 0 \\ 0 & 1+\frac{\Delta}{2} \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad (2)$$

Three parameters are thus needed to fully quantify elliptic magnifications, including the average magnification (S_{avg}), the amount of distortion (Δ), and the angle of minor axis (θ). S_{avg} is a known constant provided by user based on the calibrated instrument magnification, or optionally could be refined separately (i.e., *refineScale* in *jspr*) (Guo and Jiang, 2014). We thus will not further discuss the mean magnification here and will focus on the other two parameters, Δ and θ , for the problem of anisotropic magnification parameter determination.

2. A generalized 2-D alignment method using projection matching for determination of elliptic distortion parameters

We formulated the task of determining the anisotropic magnification parameters as a generalized 2-D alignment problem based on matching projections of 3-D reference model to experimental images. The 3-D reconstructions can be considered isotropic in magnification even when images have not been corrected for anisotropic distortion. This is a good approximation as the large number of particles in a wide range of views, including in-plane rotation angles, can effectively average out the anisotropic distortion in the 3-D model. Further, the 3-D model will become even more isotropic after imposing the symmetry, if existing. Nevertheless, the resolution of the 3-D reconstruction is inevitably reduced if the elliptic distortions of images are not properly corrected. Prior to search of the elliptic distortion parameters, other particle parameters including Euler angles, center positions, CTF parameters and mean magnifications, need to be first determined. Therefore, in our workflow, we first perform image processing tasks, including particle selection, CTF fitting, orientation/center search, 3-D reconstructions, *etc.*, as usual without determination/correction of elliptic distortion (Guo and Jiang, 2014). The 3-D models are then used to generate 2-D projections, which are computationally distorted to different combinations of α and θ parameters. Correlations of the distorted projection images with the experimental images are computed and the α - θ pair resulting in the best correlation will be accepted as the distortion parameters of the experimental images. These parameters will be used to correct anisotropic distortion in computer memory when reconstructing the 3-D models. Therefore, in our workflow, there is no need to store extra sets of micrographs or particle images with elliptic distortions corrected.

There are two parameters, α and θ , to be searched. The α and θ parameters naturally form the basis of a polar coordinate system with α along radial direction and θ along angular direction (Figure 1D). We converted these two parameters into Cartesian coordinate system ($x = r \cdot \cos\theta$, $y = r \cdot \sin\theta$), and then used x and y as direct search variables to allow a more uniform sampling of the searching space (Figure 1E). In contrast to α and θ , which are different in scale and require different step sizes for the search, x and y have the same scale and can share a common step size of searching. Recursive reduction of the search step sizes around the best search results (Figure 1E) makes the searching efficient by starting with larger step sizes (e.g. 0.01) to cover a large range (e.g., 0.05) but ending with smaller step sizes (e.g., 0.001) to reach high accuracy.

3. Implementation

The above search algorithm was implemented in C++ as a 2-D aligner *anisoscale* in our *jspr* software (Guo and Jiang, 2014) that is built based on EMAN2 library (Tang et al., 2007). It can be used alone or together with other existing 2-D aligners such as *refineOrientationCenter* and *refineMicrographDefocus* in *jspr*. The determined distortion parameters can be used to correct the distortion using the new *jspr* processor, *xform.anisoscale*, which is also implemented in C++. We typically use quadratic interpolation although both bilinear and quadratic interpolations are supported. The *xform.anisoscale* processor is internally used in both refinements and 3-D reconstruction to correct the distortion. Following the EMAN2 design, both *anisoscale* aligner and

xform.anisoscale processor are also exposed to Python scripting. The *jspr* software and user document are available on the author's Web site (<http://jiang.bio.purdue.edu/jspr>).

4. Single particle 3-D reconstructions with *jspr*

The overall process of single particle 3-D reconstructions with *jspr* has been described before (Guo and Jiang, 2014). After installation of K2 Summit camera, we have further extended it to support processing of movies. For cryo-EM data collected using DDDs, motion among raw frames was first determined using a driver script *motionCorrect.py* that runs the *dosefgpu_driftcorr* method (Li et al., 2013) in batch mode. The *dosefgpu_driftcorr* program was modified to support running average of neighboring frames (i.e., the *-nrw* option) to make the alignment trajectory smoother (supplementary Fig. 1). Both the *motionCorrect.py* script and the modified *dosefgpu_driftcorr* program are available on GitHub (<https://github.com/jianglab/motioncorr>). The resultant movie sums, which are motion-corrected but without dose-weighting, were used for particle picking using either *e2boxer.py* (Tang et al., 2007) or RELION particle picking module (Scheres, 2012). The generated coordinates files, together with the frame shift parameters and electron dose, were used to extract particles from raw movie frames with *jspr* program *batchboxer.py* while both motion correction and dose weighting were performed. Two different dose weighting strategies, mode 1 (Wang et al., 2014) and mode 2 (Grant and Grigorieff, 2015b), are incorporated into *batchboxer.py*. Similar results were obtained for these two dose-weighting modes, and the mode 1 strategy has been used for all the datasets processed in this work. The contrast transfer function (CTF) parameters were determined with *jspr* program *fitctf2.py* (Jiang et al., 2012) using the extracted particles. The whole dataset was then split into two halves, even and odd subsets, and processed independently with *jspr* as described before (Guo and Jiang, 2014; Liu et al., 2016). After reaching the convergence of Euler/center parameters, the *anisoscale* aligner was included into the iterative refinement to determine and correct the anisotropic magnification distortion, which together with Euler/center aligners (e.g., *refineOrientationCenter*) also improved the accuracy of Euler/center parameters. More aligners could be added either before or after *anisoscale* refinement, including *refineMicrographDefocus* to refine the initial defocus determined by *fitctf2.py* and also to determine per-particle defocuses, *refineAstigmatism* to determine the astigmatism aberration, *refineScale* to compensate the relative scale variance among all the 2-D images, and *refineBeamTilt* to correct potential phase errors from off-axis coma (Guo and Jiang, 2014; Liu et al., 2016). After all the refinements, the even and odd halves were pooled to generate the final 3-D reconstruction. The 0.143 cutoff of Fourier Shell Correlation (FSC) curves of the independently refined even and odd models was used to estimate the resolution (Rosenthal and Henderson, 2003).

Results and Discussion

1. The amount and angle of elliptic distortion can be reliably estimated from particle images

Multiple datasets collected using the Titan Krios TEM at different nominal magnifications were tested with the *anisoscale* aligner of *jspr* to estimate the elliptic distortion parameters. As shown in Figure 2, the amount and angle of elliptic distortion estimated using *jspr* (*jspr*,

θ_{jspr}) exhibited a pattern where a central peak was surrounded with some scattered points, when the distortion angle θ_{jspr} and the distortion amplitude γ_{jspr} are plotted in a polar coordinate system. The peak positions, representing the dominant angle and amount of the magnification distortion, were consistent among datasets that were collected within a year using the same nominal magnification of the same instrument (Fig. 2A–D for 22.5K; Fig. 2E and supplementary Fig. 4I for 18K).

To further validate these elliptic distortion parameters (γ_{jspr} , θ_{jspr}), experimental measurements of the magnification distortion for the Titan Krios TEM at Purdue University were then performed with a polycrystalline gold sample (Grant and Grigorieff, 2015c). As listed in table 1, there was ~0.5–3% elliptic distortion at nominal magnifications ranging from 14,000x to 59,000x. Comparison of the experimental magnification distortion parameters ($\gamma_{precalib}$, $\theta_{precalib}$) with the peak values of (γ_{jspr} , θ_{jspr}) revealed excellent agreement at all the tested magnifications (Fig. 2A–F; table 1). The consistent elliptic distortion parameters among multiple datasets at the same magnification of the same instrument (Fig. 2A–D; Fig. 2E and supplementary Fig. 4I), together with their agreement with the calibration measurements (table 1) indicate the reliability and robustness of the *anisoscale* aligner of *jspr* in estimating elliptic magnification distortion directly from cryo-EM particle images. This data-adaptive method of elliptic distortion estimation will make it possible to revisit and improve 3-D reconstructions of some years-old datasets without elliptic distortion calibration data.

2. Variations of the elliptic distortion resolved by *anisoscale* refinements

Furthermore, to compare the quality of 3-D reconstructions corrected using ($\gamma_{precalib}$, $\theta_{precalib}$) or (γ_{jspr} , θ_{jspr}) distortion parameters, new 3-D maps were reconstructed after replacing the (γ_{jspr} , θ_{jspr}) parameters of individual particles with ($\gamma_{precalib}$, $\theta_{precalib}$) at corresponding magnifications (table 1); FSC curves were then recalculated and compared with the original FSC curves. As shown in Figure 3, similar resolutions were achieved for 3-D reconstructions of relatively small particles (i.e., human rhinovirus of ~30 nm, and Tulane virus of ~40 nm) (Fig. 3A–B), while for particles of relatively large sizes (i.e., phage Sf6 of ~62 nm, phage W2 of ~64 nm, and T4 isometric head of ~90 nm), applying the (γ_{jspr} , θ_{jspr}) distortion parameters usually resulted in 3-D reconstructions of slightly better resolutions (Fig. 3C–E). Consistently, replacing the refined elliptic distortion parameters of individual particles with the peak values as shown in Figure 2, also resulted in 3-D reconstructions with similarly reduced quality for the larger particles (supplementary Fig. 2A–E). Unlike the calibration measurement method that assumes a constant set of distortion parameters for the entire cryo-EM dataset (Capitani et al., 2006; Grant and Grigorieff, 2015c; Zhao et al., 2015), elliptic distortion parameters of individual particles were determined with the *anisoscale* algorithm for the tested datasets. The observation of resolution reduction for 3-D reconstructions corrected using a constant value from either pre-calibration or the *jspr* computed peak value, indicates potential variations of anisotropic distortions among particles within a dataset, which has been resolved by our algorithm.

To further evaluate the spread of (γ_{jspr} , θ_{jspr}) among particles within a dataset as shown in Figure 2, different amounts of (γ_{jspr} , θ_{jspr}) “outliers” of a testing dataset of Tulane virus

recorded on photographic film (see table 2) were replaced with the peak value prior to 3-D reconstruction. Such replacement of (γ_{jspr} , θ_{jspr}) “outliers” with the peak value slightly reduced the quality of 3-D reconstruction, and the reduction of quality became more obvious when more “outliers” were replaced (supplementary Fig. 3A–B). To better understand the variations of elliptic distortion, we have investigated the particle location dependence of (γ_{jspr} , θ_{jspr}); Tulane virus particles located at the bottom and top right corners on films were found having larger γ_{jspr} , while no significant location dependence for θ_{jspr} was observed (supplementary Fig. 3C). Such location dependence of γ_{jspr} was likely due to scanner defects during digitization (Henderson et al., 2007) and it was not observed for datasets recorded with the K2 detector. All these tests suggested that the spread of the elliptic distortion parameters (γ_{jspr} , θ_{jspr}) among particles revealed by the *anisoscale* aligner indeed reflected true variable magnification distortion within a dataset to a large extent.

Since the spread was inevitably contaminated by noise caused errors given the intrinsic low signal-to-noise ratio of cryo-EM images, we also further examined the results of per-micrograph and per-particle level refinements, which can be easily specified with a “*batchsize*” option of our *anisoscale* aligner. Refinement with the *anisoscale* aligner typically started from the per-micrograph level that was followed by per-particle level refinement. Extending refinement from per-micrograph to per-particle level retained or marginally improved the reconstruction in our test cases (supplementary Fig. 4A, D and G), and per-particle level estimation and correction of elliptic distortion have been performed for all the test cases in this work. Depending on the size of particles and/or the amounts of particles in individual micrographs, per-particle level and per-micrograph level refinement and correction of the elliptic distortion may include more computing errors relative to the peak-values. Therefore, peak-value, per-micrograph and per-particle levels elliptic distortion correction can be performed separately in practice and compared for the gain or loss of 3-D reconstruction quality.

We speculated that the hysteresis of electromagnetic lenses of microscopes was possibly related to the variation of distortions among different micrographs, while potential local charging of specimens and/or other potential distortions such as scanner distortion for data collected with photographic films (Henderson et al., 2007) might further contribute to the variation of distortions among particles. However, more cases and more systematic studies will be required to fully understand the factors contributing to elliptic distortion and the variations.

3. Computational estimation and correction of the elliptic distortion significantly improves the achievable resolutions of single particle 3-D reconstructions

To evaluate the benefits of the *anisoscale* distortion estimation and correction method implemented in *jspr* for single particle 3-D reconstructions, we have tested the method with multiple datasets that led to structures of subnanometer to near-atomic resolutions without correction of the elliptic distortion. As expected, iterative estimation and correction of elliptic distortion using *anisoscale* aligner improved both the accuracy of 2-D alignments and 3-D reconstructions, and resulted in significant resolution improvements for all the

datasets containing unignorable amounts of elliptic distortions as shown in Figure 4 and (Sirohi et al., 2016).

Cryo-EM data of Tulane virus (TV, ~40 nm in diameter) (Yu et al., 2013) collected using Titan Krios at 22,500x nominal magnification on K2 detector is one case where the estimation and correction of elliptic magnification distortion using *jspr* significantly boosted the achievable resolution. As shown in Figure 4A, without including the *anisoscale* aligner, the reconstruction of TV reached only a resolution of 3.44 Å with ~14,000 particle images, while including the *anisoscale* estimation and correction improved the resolution to 2.62 Å, essentially at the Nyquist resolution (2.6 Å) of the K2 physical pixel at this magnification. At this improved resolution of ~2.6 Å, better side chain densities of amino acids were obtained for more accurate atomic model interpretation (e.g., unambiguous assignment of side chain rotamers). For example, two rotameric conformations of Isoleucine residues were recognizable in the portion of density from the 2.62 Å map shown in Figure 4B.

Another case is shown in Figure 4C, for which a significant improvement of resolution from ~4.14 Å to ~3.16 Å was obtained by performing the anisotropic scale estimation and correction using *jspr*. This cryo-EM dataset contains ~3,700 images of human rhinovirus (HRV) particles (~30 nm in diameter) (Bochkov et al., 2011) collected using the Titan Krios at 14,000x nominal magnification on the K2 Summit DDD. Such improvement in resolution was confirmed also by the obviously improved map quality. As shown in Figure 4D, compared to the original map without anisotropic magnification distortion correction, more side chain densities were resolved in the improved map at ~3.2 Å.

Since the *anisoscale* method in *jspr* does not require pre-calibration, we also revisited some of our published datasets. We first re-processed a dataset of TV collected using Titan Krios at 37,000x nominal magnification recorded on photographic films and digitized using Nikon Coolscan scanner in 2012 (Yu et al., 2013), considering the ~2.2% magnification distortion now found for the 37,000x nominal magnification of the Titan Krios at Purdue (table 1). A peak of ~2.9% elliptic distortion at ~37° was estimated by the *anisoscale* aligner (supplementary Fig. 3A), which is different from the current pre-calibration values (~2.2% at ~115°, table 1). The differences, especially the amount, are probably primarily due to the inherent variation of the microscope over time and/or after many services performed to the microscope, and also potential distortion by nonflat photographic films and/or scanner defects during digitization of the photographic films (Henderson et al., 2007). The large change of the distortion angle should be mostly due to the different orientations of the photographic films and the K2 detector. Both our previous refinement and new refinement without considering the elliptic distortion using *jspr* led to a 3-D reconstruction of ~6.1–6.3 Å (Yu et al., 2013), while further refinement using the *anisoscale* aligner improved the reconstruction markedly to ~4.2 Å resolution (Fig. 4E), at which individual strands of β sheet became resolved (Fig. 4F). Introduction of the *anisoscale* aligner to the single particle 3-D reconstruction of a dataset of Porcine circovirus type 2 (PCV2, ~19 nm in diameter) collected using Titan Krios at 59,000x nominal magnification on photographic films in 2011, however, resulted in much smaller improvement in resolutions (from ~3.0 to 2.9 Å) (Liu et al., 2016), presumably due to the small amount of distortion at this high magnification of 59,000x (table 1) and the small size of the PCV2 particle.

One more case on particles of much larger size is shown in Figure 4G/H. Refinement of a dataset of a bacteriophage T4 mutant with an isometric head (~90 nm in diameter) (Olson et al., 2001) was limited to ~10 Å with ~18,000 particle images collected using Titan Krios at 18,000x nominal magnification on K2 detector. The inclusion of *anisoscale* aligner in *jspr* refinement to estimate and correct the elliptic distortion drastically improved the attainable resolution of this bacteriophage T4 dataset from subnanometer to near-atomic resolution (3.27 Å, Fig. 4G/H). A 3-D reconstruction using the finally refined parameters without elliptic distortion correction reached a resolution of ~7.16 Å (Fig. 4G) that was also notably better than original refinement without considering the elliptic distortion (~10 Å). This indicated that the correction of elliptic distortion also improved the accuracy of Euler/Center parameters, which was consistent among our tested cases.

In addition to our own datasets collected at Purdue, we also tested our method with a public dataset for the 2.8 Å structure of *Thermoplasma acidophilum* 20S proteasome (~18 nm in the long dimension) (Campbell et al., 2015) deposited in the Electron Microscopy Pilot Image Archive of EMDataBank (EMPIAR-10025). Although this dataset of 20S proteasome was collected using the same nominal magnification (22,500x) of Titan Krios on K2 detector as some of our datasets collected at Purdue University, little anisotropic magnification distortion was observed as revealed by the per-micrograph level refinement with *anisoscale* (supplementary Fig. 5A), suggesting that the Titan Krios at The Scripps Research Institute did not have significant elliptic lens distortion at 22,500x nominal magnification; 3-D reconstructions of similar resolutions (~2.65 Å) were obtained with/without this refinement (supplementary Fig. 5C). Further refinement with *anisoscale* at per-particle level introduced a larger spread of (θ_{jspr}) (supplementary Fig. 5B) potentially due to increased randomness of elliptic distortion among individual particles and more computing errors of *anisoscale* with individual images of low SNRs, while resulting in a 3-D reconstruction at a similar resolution (supplementary Fig. 5C). This indicated that the potential gain from resolving elliptic distortion variations among particles (e.g., due to local charging of specimens) by *anisoscale* was balanced by loss due to computing errors for the small 20S proteasome particles, and there was no net gain or loss of reconstruction quality.

4. Correlation of resolution limit and resolution improvement with elliptic distortion

Additional datasets have also been tested, and the results of all tested cases were summarized in table 2. A linear correlation between the product of particle diameter and the dominant elliptic distortion amplitude (D^*) and the achievable resolutions without the distortion correction (resolution_limit) was observed based on these tested cases (Fig. 5A). As expected, cryo-EM datasets with larger D^* are more severely affected and limited to lower resolutions. A published rotavirus dataset with elliptic distortion (Grant and Grigorieff, 2015a) also followed this trend. This plot can be used as a guide to assess if elliptic distortion is a major resolution limiting factor for cryo-EM datasets. Resolution improvements ($res(-anisoscale)/res(+anisoscale)$, table 2) ranging from ~1.3 to 3.1 folds were obtained for the tested datasets collected at nominal magnifications ranging from 14,000x to 37,000x bearing ~2–3% elliptic distortion. Such resolution improvements were also found strongly correlated with D^* (Fig. 5B). In general, datasets with larger D^* benefits more from the elliptic distortion correction. The aforementioned rotavirus dataset

(Grant and Grigorieff, 2015a) also agreed well with the approximated trend line that can serve as a guideline to predict the potential benefits of elliptic distortion correction.

Conclusions

Elliptic distortion of TEM images is introduced mainly by projector lenses (e.g., the P2 lens for the Titan Krios at Purdue according to preliminary tests using cross-grating grids by previous facility manager Dr. Agustin Avila-Sakar, unpublished), and potentially further modulated by other issues in imaging such as sample charging. It can be a resolution limiting factor of single particle 3-D reconstructions in a large range of resolutions (Fig. 5). Correction of the elliptic distortion is essential for obtaining 3-D reconstructions of large macromolecules such as most viral particles at sub-3 Å and better resolutions, while for smaller protein complexes, it could become a problem when targeting at higher resolutions (e.g., beyond 2 Å). In this work, we have developed a generic algorithm for the elliptic anisotropic magnification distortion estimation and correction in single particle 3-D reconstructions. This method brings multiple benefits, including 1) it eliminates the need of pre-calibration using standard specimens; 2) it is data-adaptive and therefore is generally applicable to any single particle cryo-EM datasets irrespective of facility settings; 3) The per-micrograph and/or per-particle level determination of anisotropic magnification distortion in our algorithm alleviates the concern about the potentially variable anisotropic magnification distortions caused by charging of the samples and numerous changes of currents of all lenses during data collection spanning hours to days; 4) In our method, the raw images were left unmodified and the distortion correction is performed during the 3-D reconstruction step, and therefore there is no need for extra space to store distortion corrected micrographs.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported in part by NIH grant (1R01AI111095). Cryo-EM images were taken at the Purdue Cryo-EM Facility (<http://cryoem.bio.purdue.edu>).

References

- Bai XC, Rajendra E, Yang G, Shi Y, Scheres SH. Sampling the conformational space of the catalytic subunit of human gamma-secretase. *Elife*. 2015; 4:e11182. [PubMed: 26623517]
- Banerjee S, Bartesaghi A, Merk A, Rao P, Bulfer SL, Yan Y, Green N, Mroczkowski B, Neitz RJ, Wipf P, Falconieri V, Deshaies RJ, Milne JL, Huryn D, Arkin M, Subramaniam S. 2.3 Å resolution cryo-EM structure of human p97 and mechanism of allosteric inhibition. *Science*. 2016; 351:871–875. [PubMed: 26822609]
- Bartesaghi A, Merk A, Banerjee S, Matthies D, Wu X, Milne JL, Subramaniam S. 2.2 Å resolution cryo-EM structure of beta-galactosidase in complex with a cell-permeant inhibitor. *Science*. 2015; 348:1147–1151. [PubMed: 25953817]
- Bochkov YA, Palmenberg AC, Lee WM, Rathe JA, Amineva SP, Sun X, Pasic TR, Jarjour NN, Liggett SB, Gern JE. Molecular modeling, organ culture and reverse genetics for a newly identified human rhinovirus C. *Nat Med*. 2011; 17:627–U152. [PubMed: 21483405]

- Campbell MG, Veesler D, Cheng A, Potter CS, Carragher B. 2.8 Å resolution reconstruction of the *Thermoplasma acidophilum* 20S proteasome using cryo-electron microscopy. *Elife*. 2015; 4:e06380.
- Capitani GC, Oleynikov P, Hovmöller S, Mellini M. A practical method to detect and correct for lens distortion in the TEM. *Ultramicroscopy*. 2006; 106:66–74. [PubMed: 16046067]
- Clemens DL, Ge P, Lee BY, Horwitz MA, Zhou ZH. Atomic structure of T6SS reveals interlaced array essential to function. *Cell*. 2015; 160:940–951. [PubMed: 25723168]
- Fokine A, Battisti AJ, Kostyuchenko VA, Black LW, Rossmann MG. Cryo-EM structure of a bacteriophage T4 gp24 bypass mutant: the evolution of pentameric vertex proteins in icosahedral viruses. *J Struct Biol*. 2006; 154:255–259. [PubMed: 16530424]
- Grant T, Grigorieff N. Automatic estimation and correction of anisotropic magnification distortion in electron microscopes. *J Struct Biol*. 2015a; 192:204–208. [PubMed: 26278979]
- Grant T, Grigorieff N. Measuring the optimal exposure for single particle cryo-EM using a 2.6 Å reconstruction of rotavirus VP6. *Elife*. 2015b; 4:e06980. [PubMed: 26023829]
- Guo F, Jiang W. Single particle cryo-electron microscopy and 3-D reconstruction of viruses. *Methods Mol Biol*. 2014; 1117:401–443. [PubMed: 24357374]
- Guo F, Liu Z, Fang PA, Zhang Q, Wright ET, Wu W, Zhang C, Vago F, Ren Y, Jakana J, Chiu W, Serwer P, Jiang W. Capsid expansion mechanism of bacteriophage T7 revealed by multistate atomic models derived from cryo-EM reconstructions. *P Natl Acad Sci USA*. 2014; 111:E4606–4614.
- Henderson R, Cattermole D, McMullan G, Scotcher S, Fordham M, Amos WB, Faruqi AR. Digitisation of electron microscope films: Six useful tests applied to three film scanners. *Ultramicroscopy*. 2007; 107:73–80. [PubMed: 16872749]
- Hou V. A DigitalMicrograph™ Script to Characterize Elliptical Distortion of Electron Diffraction Patterns in TEM. *Microscopy and Microanalysis*. 2008; 14:1124–1125.
- Jiang J, Pentelute BL, Collier RJ, Zhou ZH. Atomic structure of anthrax protective antigen pore elucidates toxin translocation. *Nature*. 2015; 521:545–549. [PubMed: 25778700]
- Jiang W, Chiu W. Cryoelectron microscopy of icosahedral virus particles. *Methods Mol Biol*. 2007; 369:345–363. [PubMed: 17656759]
- Jiang W, Guo F, Liu Z. A graph theory method for determination of cryo-EM image focuses. *J Struct Biol*. 2012; 180:343–351. [PubMed: 22842112]
- Li XM, Mooney P, Zheng S, Booth CR, Braunfeld MB, Gubbens S, Agard DA, Cheng YF. Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nat Methods*. 2013; 10:584–590. [PubMed: 23644547]
- Liao M, Cao E, Julius D, Cheng Y. Structure of the TRPV1 ion channel determined by electron cryo-microscopy. *Nature*. 2013; 504:107–112. [PubMed: 24305160]
- Liu H, Jin L, Koh SB, Atanasov I, Schein S, Wu L, Zhou ZH. Atomic structure of human adenovirus by cryo-EM reveals interactions among protein networks. *Science*. 2010; 329:1038–1043. [PubMed: 20798312]
- Liu Z, Guo F, Wang F, Li TC, Jiang W. 2.9 Å Resolution Cryo-EM 3D Reconstruction of Close-Packed Virus Particles. *Structure*. 2016; 24:319–328. [PubMed: 26777413]
- Ludtke SJ, Baldwin PR, Chiu W. EMAN: Semiautomated software for high-resolution single-particle reconstructions. *J Struct Biol*. 1999; 128:82–97. [PubMed: 10600563]
- Mugnaioli E, Capitani G, Nieto F, Mellini M. Accurate and precise lattice parameters by selected-area electron diffraction in the transmission electron microscope. *Am Mineral*. 2009; 94:793–800.
- 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]
- Scheres SH. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J Struct Biol*. 2012; 180:519–530. [PubMed: 23000701]
- Tang G, Peng L, Baldwin PR, Mann DS, Jiang W, Rees I, Ludtke SJ. EMAN2: an extensible image processing suite for electron microscopy. *J Struct Biol*. 2007; 157:38–46. [PubMed: 16859925]

- Wang Z, Hryc CF, Bammes B, Afonine PV, Jakana J, Chen DH, Liu XG, Baker ML, Kao C, Ludtke SJ, Schmid MF, Adams PD, Chiu W. An atomic model of brome mosaic virus using direct electron detection and real-space optimization. *Nat Commun.* 2014; 5:5808. [PubMed: 25510333]
- Yu G, Zhang D, Guo F, Tan M, Jiang X, Jiang W. Cryo-EM structure of a novel calicivirus, Tulane virus. *PLoS One.* 2013; 8:e59817. [PubMed: 23533651]
- Yu X, Jin L, Zhou ZH. 3.88 Å structure of cytoplasmic polyhedrosis virus by cryo-electron microscopy. *Nature.* 2008; 453:415–419. [PubMed: 18449192]
- Zhang X, Jin L, Fang Q, Hui WH, Zhou ZH. 3.3 angstrom Cryo-EM Structure of a Nonenveloped Virus Reveals a Priming Mechanism for Cell Entry. *Cell.* 2010; 141:472–482. [PubMed: 20398923]
- Zhang X, Settembre E, Xu C, Dormitzer PR, Bellamy R, Harrison SC, Grigorieff N. Near-atomic resolution using electron cryomicroscopy and single-particle reconstruction. *P Natl Acad Sci USA.* 2008; 105:1867–1872.
- Zhao J, Brubaker MA, Benlekhir S, Rubinstein JL. Description and comparison of algorithms for correcting anisotropic magnification in cryo-EM images. *J Struct Biol.* 2015; 192:209–215. [PubMed: 26087140]

Figure 1. Illustration of elliptic anisotropic magnification distortion

(A) Isotropic magnification in the absence of elliptic distortion. (B) Elliptic distortion with the major axis along Y-axis and minor axis along X-axis. (C) A generalized form of elliptic distortion when the minor axis is at an arbitrary angle θ . (D) Uneven sampling of Δ and θ in a polar coordinate system with Δ along radial direction and θ along angular direction. Sampling of the space is finer at smaller Δ , while much coarser at larger Δ . (E) Uniform sampling of the search space of elliptic distortion parameters by converting into Cartesian coordinate system with $x = \Delta \cos\theta$, $y = \Delta \sin\theta$ as the search variables. Recursive decrease of the step size was performed to find the final solution (Δ) at high accuracy.

Figure 2. Elliptic distortion parameters (j_{spr} , θ_{jspr}) determined using *anisoscale* aligner in *jspr* Multiple datasets collected at different magnifications using the Titan Krios at Purdue recorded with K2 electron detector were tested. (A–D) Four datasets, Tulane virus (TV), enterovirus D68 (EV-D68), phage ST1 and phage W2, were imaged at the same nominal magnification of 22,500×. (E) The phage T4 dataset was imaged at 18,000× nominal magnification. (F) The human rhinovirus (HRV) dataset was imaged at 14,000× nominal magnification. j_{spr} and θ_{jspr} were plotted in a polar coordinate system with j_{spr} along radial direction and θ_{jspr} along angular direction. The peak positions indicating the dominant elliptic distortion parameters were labeled accordingly.

Figure 3. Resolution comparison of 3-D reconstructions with elliptic distortion corrected based on (θ_{precalib} , θ_{jspr}) and (θ_{precalib} , θ_{jspr})

FSC curves from five datasets of (A) human rhinovirus (HRV), (B) Tulane virus (TV), (C) phage Sf6, (D) phage W2 and (E) phage T4, covering particles of different sizes (~30–90nm), were plotted. The blue curves represent FSCs calculated based on maps with elliptic distortion corrected with (θ_{jspr} , θ_{jspr}), and the green ones are from maps with elliptic distortion corrected with (θ_{precalib} , θ_{precalib}). Corresponding resolutions at 0.143 cutoff of FSC were labeled. The data collection session of phage W2 was interrupted for re-alignments of the microscope due to beam instability; The large discrepancy of the two FSC curves in (D) likely also reflects the extra variations of instrument conditions.

Figure 4. Improved quality of 3-D reconstructions after *anisotropy* estimation and correction of elliptic distortion

Results from four datasets covering different resolutions ranging from subnanometer to near-atomic resolution are displayed. Shown are four datasets: (A, B) Tulane virus (TV) (K2 detector), (C, D) human rhinovirus (HRV), (E, F) TV (previously published dataset imaged on film (Yu et al., 2013)), and (G, H) phage T4. FSC curves from 3-D reconstructions with (blue)/without (green) including *anisotropy* aligner in iterative refinement were shown (A, C, and E). (B, D, and F) illustrate the map quality improvement; Panels in the left show the maps from refinements without *anisotropy* aligner, and those in the right are from refinements with the *anisotropy* aligner. Arrows in (B) point to two Ile residues with different side chain rotamer conformations. (G and H) show the FSC curves, and the maps of phage T4 final reconstruction with/without performing the elliptic distortion correction, respectively.

Figure 5. Resolution limits, improvements and their dependence on particle size and the amount of elliptic distortion

Resolution limits (A) and improvements (B) as function of the product of particle size and the amount of elliptic distortion. The trend lines were fitted using the data points labeled with blue dots representing the datasets collected using Titan Krios at Purdue listed in table 2. Resolution_limit trend line: $y = 0.3485x + 0.7832$ ($R^2 = 0.8407$); The improvements trend line: $y = 0.0951x + 0.4728$ ($R^2 = 0.9549$). The data point represented by with diamond-shaped symbols is from the Rotavirus case reported in (Grant and Grigorieff, 2015a).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1
Experimental measurement of anisotropic magnification distortion of the Titan Krios microscope at Purdue University
Images of a polycrystalline gold sample were taken with a Gatan K2 Summit direct electron detector under different nominal magnifications, and the anisotropic magnification distortion parameters were determined using *mag_distortion_estimate* program (Grant and Grigorieff, 2015c).

nominal mag	α (degree)	minor axis scale factor	major axis scale factor	%distortion	θ_{preclib} *	θ_{preclib} *
14,000	31.3	0.986	1.014	2.87	121.3	0.028
18,000	28.1	0.985	1.015	3.02	118.1	0.030
22,500	28.2	0.987	1.013	2.72	118.2	0.026
29,000	24.8	0.987	1.013	2.57	114.8	0.026
37,000	24.9	0.989	1.011	2.16	114.9	0.022
47,000	34.8	0.992	1.008	1.61	124.8	0.016
59,000	60.2	0.998	1.002	0.55	150.2	0.004

* The determined parameters were converted to corresponding parameters in *jspr* as described in Methods section, named θ_{preclib} and θ_{preclib} .

Table 2
Resolution improvements of 3-D reconstructions by anisotropic magnification distortion estimation and correction with the *anisoscale* aligner of *jspr*

Multiple datasets collected using Titan Krios microscope and K2 camera at Purdue with a dose rate of 8 electrons/physical pixel/second and frame rate of 5 frames/second were tested. The datasets were Tulane virus (TV), enterovirus D68 (EV-D68), phage ST1, Sf6, W2 and T4, and human rhinovirus (HRV). Three previously published datasets^{1–3} were also tested.

mag	detector	sample	diameter /nm	total mass/MDa	res/Å (<i>-anisoscale</i>)	res/Å (<i>+anisoscale</i>)	improvement ⁴	#ptcls
14,000	K2	HRV	30	5.8	4.14	3.16	1.3	3,739
18,000	K2	phage T4 ⁵	90	42.9	~10	3.27	3.1	17,954
18,000	K2	phage Sf6	62	19.5	6.05	2.89	2.1	68,237
22,500	K2	EV-D68	32	5.0	3.20	2.42	1.3	11,344
22,500	K2	phage ST1	30	6.0	3.20	2.44	1.3	32,120
22,500	K2	TV	40	10.4	3.44	2.62	1.3	14,154
22,500	K2	phage W2	64	20.5	8.23	3.59	2.3	8,663
37,000	film	TV ¹	40	10.4	6.13	4.17	1.5	4,278
59,000	film	PCV2 ²	19	1.7	~3	2.92	1.03	55,178
22,500	K2	20S proteasome ³	18	0.8	2.65	2.65	1	49,954

¹Dataset reported in (Yu et al., 2013).
²Dataset reported in (Liu et al., 2016).
³Dataset reported in (Campbell et al., 2015) that was imaged using the Titan Krios at The Scripps Research Institute.
⁴The “improvement” is computed as $\text{res}(-\text{anisoscale})/\text{res}(+\text{anisoscale})$.
⁵A T4 mutant with icosahedral head.