Fast rotational matching of single-particle images

Yao Cong [a], Wen Jiang [b], Stefan Birmanns [a], Z. Hong Zhou [c], Wah Chiu [b], Willy Wriggers [a,*]

[a] *School of Health Information Sciences and Institute of Molecular Medicine, University of Texas Health Science Center at Houston,
7000 Fannin St., Suite 600, Houston, TX 77030, USA*
[b] *Verna and Marrs McLean Department of Biochemistry and Molecular Biology, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA*
[c] *Department of Pathology and Laboratory Medicine, University of Texas-Houston Medical School, TX 77030, USA*

Abstract

The presence of noise and absence of contrast in electron micrographs lead to a reduced resolution of the final 3D reconstruction, due to the inherent limitations of single-particle image alignment. The fast rotational matching (FRM) algorithm was introduced recently for an accurate alignment of 2D images under such challenging conditions. Here, we implemented this algorithm for the first time in a standard 3D reconstruction package used in electron microscopy. This allowed us to carry out exhaustive tests of the robustness and reliability in iterative orientation determination, classification, and 3D reconstruction on simulated and experimental image data. A classification test on GroEL chaperonin images demonstrates that FRM assigns up to 13% more images to their correct reference orientation, compared to the classical self-correlation function method. Moreover, at sub-nanometer resolution, GroEL and rice dwarf virus reconstructions exhibit a remarkable resolution gain of 10–20% that is attributed to the novel image alignment kernel.
© 2005 Elsevier Inc. All rights reserved.

Keywords: Electron microscopy; Single-particle analysis; 2D alignment; 3D reconstruction; Image processing; Fast rotational matching; GroEL; RDV

1. Introduction

In single-particle analysis of electron microscopy (EM) images, a reference-based refinement strategy typically involves the following three steps (Baker and Cheng, 1996; Baker et al., 1999; Frank et al., 1996; Ludtke et al., 1999, 2004; van Heel et al., 1996): (1) classification, which determines the two out-of-plane rotational parameters among the five parameters associated with each raw image; an image is compared to a set of reference projection images generated from a 3D model and then assigned to the most similar projection class. (2) Class averaging, which determines the remaining three in-plane parameters; here, the particles within a class are aligned to the reference image and then averaged to generate a new class average to enhance the signal-to-noise ratio (SNR). (3) Reconstruction, where the class averages with assigned Euler angles are used to construct a new 3D model for the next round of refinement.

In the classification and class averaging, where the five rigid body parameters of each raw particle image are assigned, the 2D image alignment algorithm determines the accuracy of orientation and center assignment, which ultimately affects the quality of the 3D reconstruction. Therefore, the 2D alignment remains as one of the essential algorithmic performance bottlenecks in single-particle reconstruction. In contrast to the experimental limitations such as heterogeneity of the data and instrument parameters (Baldwin and Penczek, 2005; Chiu et al., 2005; Cong et al., 2003; Frank, 1996; Joyeux and Penczek, 2002), limitations in 2D alignment can be remedied computationally in the post-processing stage of collected images.

In previous work, the fast rotational matching (FRM) algorithm was introduced, where the in-plane alignment was achieved by a systematic search of three parameters (Cong et al., 2003): the 2D fast Fourier transform (FFT) was used to accelerate a 2D rotational search while the remaining single translational parameter was obtained by

---
* Corresponding author. Fax: +1 713 500 3907.
E-mail address: wriggers@biomachina.org (W. Wriggers).

exhaustive scanning within a limited range. Based on 2D images it was demonstrated that FRM combines the accuracy of polar coordinate sampling (RPC) (Joyeux and Penczek, 2002; Penczek et al., 1992) with the efficiency of self-correlation function (SCF) alignment (Frank et al., 1978; Schatz and van Heel, 1992; van Heel et al., 1992). However, in the original work the FRM algorithm was only evaluated in 2D on simulated EM images. Here, we have evaluated FRM performance for the first time on actual experimental images and in the context of a complete iterative 3D reconstruction project. The evaluation on 3D data presented in the following offers novel insights into the dependence of EM map resolution on the 2D image alignment quality.

In the open source package EMAN for single-particle reconstruction from transmission electron micrographs (Ludtke et al., 1999, 2001, 2004), the SCF method is the default alignment kernel in the classification and the class averaging procedures. The SCF method is computationally efficient; however, it suffers from the intrinsic sensitivity to noise, especially when dealing with data at low SNR (Cong et al., 2003; Joyeux and Penczek, 2002). To improve the precision of this alignment routine, a local refinement step with a simplex minimizer was originally added after the SCF alignment routine, producing alignments with sub-pixel and sub-degree precision (Ludtke et al., 2004). In this work, we fully integrated FRM as an alternative alignment function in EMAN, specifically, in the classification program *classesbymra* and the class averaging program *classalign2*. This enables us, for the first time, to utilize FRM as the alignment kernel as part of the 3D reconstruction steps.

Below, a classification accuracy test (to determine the two out-of-plane parameters) is conducted on a simulated GroEL data set containing 930 randomly rotated, translated, noise-corrupted images at all possible projection directions. Because SCF is the default alignment algorithm in EMAN, we also provide the performance results of SCF as a performance standard. Besides, to evaluate the overall accuracy of FRM and its robustness to noise, reconstructions were performed on simulated data sets of GroEL at different noise levels. In these tests, reconstructions with and without the additional sub-pixel refinement are conducted to evaluate the performance of pure alignment algorithms as well as the effect of the local refinement in the 3D reconstruction process. Finally, to validate the reliability of FRM in dealing with realistic noisy data sets, the state-of-the-art 3D reconstruction of a rice dwarf virus (RDV) is presented. RDV was chosen because a 6.8 Å cryogenic EM map (Zhou et al., 2001) and an atomic structure (Nakagawa et al., 2003) were available that enabled a detailed validation of the precision of the 3D reconstruction process.

2. Methods

2.1. Fast rotational matching

We provide a brief summary of FRM, which has been described in more detail elsewhere (Cong et al., 2003; Kovacs et al., 2003). In 2D we express the relative positions and rotations of two objects f and g (to be matched) by two FFT-accelerated angular parameters and one remaining linear parameter. We rotate both objects about their respective center of mass while translating one of them along the positive x-axis by a distance ρ (illustrated in Fig. 1 of Cong et al., 2003). We resample the density to polar coordinates, and the density functions become functions of the radius r and the polar angle β (Cong et al., 2003). Then, we expand a target density $f(r, \beta)$ and a probe density $g(r, \beta)$ into Fourier space,

$$f(r, \beta) = \sum_{m=1-B}^{B-1} \hat{f}_m(r) e^{im\beta} \tag{1}$$

(similar for g), where $\hat{f}_m$ is the Fourier coefficient and the number of B (bandwidth) is chosen according to the desired angular sampling rate. Using the expansions of the rotated and translated objects, we arrive eventually at an expression for the FFT of the correlation function C as a function of reciprocal angles m and n, and the linear scan range ρ:

$$\hat{c}(m, n; \rho) = \text{FFT}_{2D} C = 2\pi \int_0^\infty (\hat{h}_{r,\rho}^n)_m \cdot \overline{\hat{f}_m(r)} \cdot r \, dr, \tag{2}$$

where $\overline{\hat{f}_m(r)}$ is the complex conjugate of the Fourier coefficient of object f, $(\hat{h}_{r,\rho}^n)_m$ is the factor containing all the information about object g, which can be evaluated numerically. The inverse FFT then yields C and the corresponding rotational parameters for each value of ρ. Then, a peak search strategy is carried out to determine the maximum correlation value and the two corresponding rotational angles. We repeat this procedure for each ρ within a certain range. In practice, for alignment of images of roughly centered particles, the maximum ρ will be small which allows for an efficient scan.

2.2. Classification test image generation

A 3D electron density map of GroEL was generated from PDB entry 1OEL after Gaussian low-pass filtering to 5 Å using the EMAN program *pdb2mrc* (Jiang et al., 2001; Ludtke et al., 1999, 2004). The map size was $128 \times 128 \times 128$ voxels sampled at 1.9 Å/pixel. Ninety-three projections of this reference map covering all projection directions at an interval of 4° were generated using real-space projection with tri-linear interpolation by the EMAN program *project3d*. Then, each of the 93 projections was randomly rotated and translated 10 times, thus 930 raw images were generated. Besides, to simulate the particle flip in real experimental data, half of the 930 raw images were flipped. In addition, we added Gaussian white noise (using EMAN program *proc2d*) to the data, generating four test data sets with SNR levels of 0.15, 0.06, 0.04, and 0.03. The SNR is defined as $\sigma_{signal}^2 / \sigma_{noise}^2$, following Joyeux and Penczek (2002) and Cong et al. (2003). These randomly rotated, translated, flipped, and noise-corrupted images were then used in the classification test.

2.3. Reconstruction test image generation

To verify the overall performance of FRM alignment in 3D reconstruction, projection/reconstruction tests at different noise levels were performed on simulated images, again using GroEL (PDB entry 1OEL) as the test molecule. Here, the simulated 5 Å structure was used to generate 4443 projections covering all projection directions. Each image was randomly rotated and translated and Gaussian white noise was applied, generating three testing data sets with SNR at 0.15, 0.06, and 0.04, respectively. The image size was 128 × 128 pixels with 1.9 Å/pixel. Usually, the 3D reconstruction is performed in two stages: initial model generation and model refinement. Since we concentrate on the alignment kernel performance, only the second stage was conducted. The initial 3D model for the model refinement was generated from the same atomic structure after a Gaussian low-pass filtering to ∼21 Å resolution using the EMAN program *pdb2mrc* (Jiang et al., 2001; Ludtke et al., 1999, 2004). For each noise level, a 3D reconstruction with five rounds of iteration was conducted using FRM and SCF alignment kernel, respectively, with 1.4° angular sampling steps for projections. No filtering was applied in the final 3D reconstructions in all our GroEL reconstruction tests.

2.4. 3D reconstruction of icosahedral particles using EMAN

EMAN was originally designed for image processing and 3D reconstruction of single particles of low or no symmetry. As a result, support for large viruses of icosahedral symmetry was less robust in both algorithms themselves and algorithm implementations. Improvements in EMAN have been made for processing large icosahedral viruses. The overall logistics and programs of EMAN, such as phase flipping in the preprocessing step, classification by projection matching, CTF amplitude correction during class averaging, and 3D reconstruction from class averages using direct Fourier inversion, are maintained. The improvements are mostly underlying modifications to the EMAN refinement command *refine* and the lower level programs indirectly called by *refine*, and are thus largely transparent to the users. With these improvements, EMAN now can process large icosahedral virus images as well as other small lower symmetry particle images with essentially the same user interface. The modifications have been validated with simulated data and experimental data of Rice Dwarf Virus. The improvements will be described in more detail in the EMAN documentation.

3. Results and discussion

In the following, we implemented FRM into EMAN and carried out controlled classification accuracy tests as well as 3D reconstructions at different noise levels using simulated GroEL images. In addition, a reconstruction from experimental RDV data is presented to validate the reliability of FRM in dealing with real electron micrographs.

We note that a detailed efficiency comparison among FRM, SCF, and RPC in 2D has already been presented elsewhere (Cong et al., 2003). For a full classification, the total time spent on FRM is simply the 2D image alignment time described elsewhere (Cong et al., 2003) multiplied by the number of projection images. Also, our previous 2D accuracy tests (Cong et al., 2003) showed that FRM and RPC accuracies are comparable. Therefore, we have focused in this work on the accuracy comparison of FRM with SCF only, and on the way the overall quality of the 3D reconstruction is affected by these two algorithms.

3.1. Classification accuracy

To evaluate the precision of the alignment kernels, we carried out image classification tests at different noise levels for both FRM and SCF. As a test case we chose GroEL, which is a double-ring complex of two heptameric rings stacked back-to-back, practically exhibiting D7 symmetry (there are negligible deviations from 7-fold symmetry in the known atomic structure that are irrelevant for the low resolution tests conducted here). Four data sets with SNR at 0.15, 0.06, 0.04, and 0.03 were generated as described in Section 2. Representative raw images at different noise levels are shown in Fig. 1. In the tests within this section we used a 1.4° angular sampling for FRM and a 1.2° angular sampling for SCF (the default value in EMAN for this image size).

3.1.1. Alignment kernel performance
The performance of the two pure alignment algorithms (without any additional sub-pixel refinement) in terms of
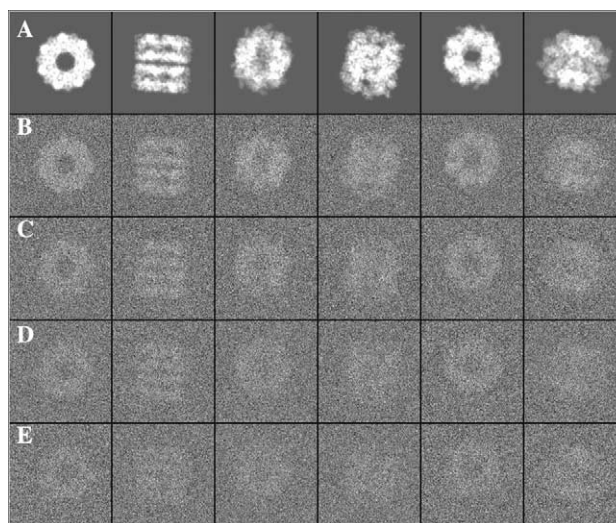

Fig. 1. Representative particles used in the classification tests. (A) Six random projections of GroEL from PDB entry 1OEL after Gaussian low-pass filtering to 5 Å resolution (cf. Section 2). (B–E) Projections in (A) corrupted with Gaussian noise with SNR of 0.15, 0.06, 0.04, and 0.03, respectively. Image size is 128 × 128 pixels with 1.9 Å/pixel.

correctly classified images is shown in Table 1. FRM improves over SCF especially in cases of low SNR. For instance, when SNR decreases to 0.03, ~12% more particles are correctly classified by FRM compared to SCF. The classification error distributions at SNR 0.03 (Figs. 2A and B) demonstrate that the FRM classification error is more localized within the 21 nearest-neighbor classes, whereas SCF exhibits a wider distribution over 67 out of a total of 93 classes. Only at a relatively high SNR level of 0.15 did SCF perform slightly better than FRM (2% more images correctly classified). This phenomenon may be attributable to a slightly better angular sampling preci-

sion of SCF (1.2°) than FRM (1.4°) and a better performance of SCF at such high SNR levels (Cong et al., 2003). However, electron micrographs are usually noisier, so in typical EM applications FRM would perform better.

This performance advantage of FRM at low SNR can be explained by the algorithmic details. First, in the SCF method, the rotational angle is determined through calculating the angle between the SCFs of a pair of images. Here, the SCF is the inverse Fourier transform of the amplitude spectrum of an image (Schatz and van Heel, 1992; van Heel et al., 1992, Eq. (5)), therefore, only the amplitude is considered while the important phase

Table 1
Classification accuracy achieved by FRM and SCF at different noise levels with and without the additional sub-pixel refinement[a]

SNR	Correctly classified image numbers			
	0.15	0.06	0.04	0.03
FRM	759 (82%)[b]	646 (69%)	536 (58%)	460 (49%)
SCF	777 (84%)	570 (61%)	455 (49%)	346 (37%)
FRM + refine	892 (96%)	768 (83%)	696 (75%)	603 (65%)
SCF + refine	872 (94%)	742 (80%)	626 (67%)	481 (52%)

[a] Each data set contains 930 raw images and 93 projections.
[b] The value in the parentheses indicates the percentage of the correctly classified images among the total 930 images.
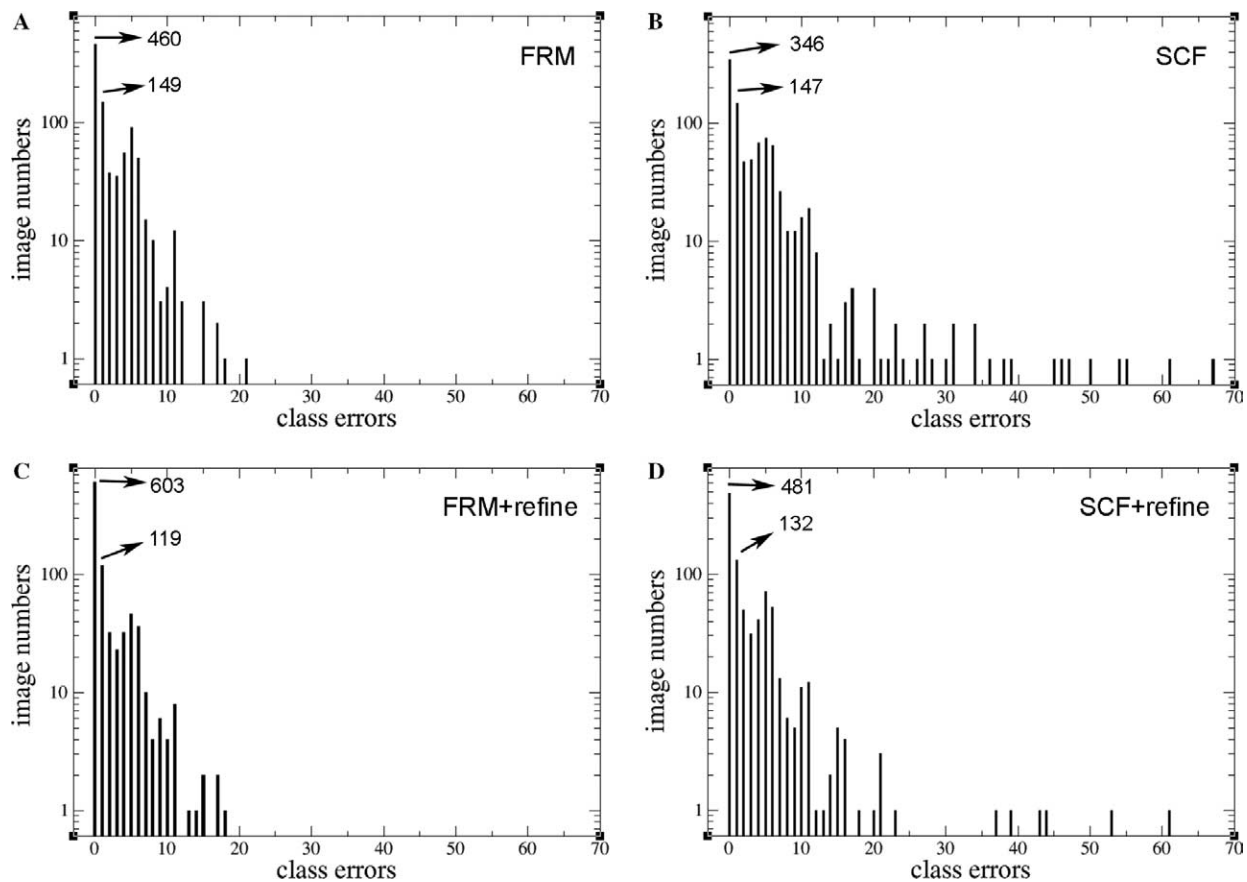

Fig. 2. Illustration of classification error distributions at SNR 0.03 with and without sub-pixel local refinement. (A) Pure FRM alignment kernel classification error distribution. Four hundred and sixty images are assigned to the correct classes, while 149 images are assigned to the nearest neighbor classes, with 4° interval between each class. (B) Pure SCF alignment kernel classification error distribution, which shows a much wider class error distribution of SCF (67) than the pure FRM (21). (C and D) FRM + refine/SCF + refine classification error distribution. Even with the local refinement step, SCF + refine still shows much wider class error distribution (61) than FRM + refine (18).

information is ignored. But in FRM both the amplitude and the phase information are fully employed in the calculation. Moreover, since the FFT operates on cyclical angle variables in FRM, there is no boundary problem typically associated with the use of FFT in linear variables such as in SCF. Finally, FRM yields the rotational and translational parameters simultaneously, avoiding any rotational interpolation error associated with SCF.

3.1.2. Effect of local refinement

To test the effect of the additional sub-pixel local refinement on the classification accuracy, we repeated the above test with the local refinement option turned on. The results listed in the lower section of Table 1 clearly indicate the advantage of using this refinement for both FRM and SCF. For instance, the correctly classified image numbers increase by more than 10% for both methods at different noise levels.

While the local refinement step is beneficial to both methods, FRM + refine still yields a better classification performance. For example, at SNR 0.03 FRM + refine successfully assigns 122 more raw images (13% of a total of 930) to the correct classes compared to SCF + refine. Clearly, if the initial alignment is unsuccessful, a local refinement has limited chance to correct it. This can also be validated by the classification error distributions illustrated in Fig. 2. The pure SCF classifications (Fig. 2B) span a wide error range of up to 67 classes; and the local refinement decreases this error range only slightly to 61 classes (Fig. 2D).

In summary, the above classification tests demonstrate that FRM can considerably improve the classification

accuracy, which makes better use of available raw images and promises to yield higher quality 3D reconstructions.

3.2. 3D reconstruction performance

In this section, various reconstructions of simulated GroEL images at three different noise levels are presented to validate the robustness of FRM in the overall 3D reconstruction procedure. The corresponding SCF reconstructions are presented as a control.

3.2.1. Alignment kernel performance

In Figs. 3A–C, we illustrate the resolution assessments of reconstructions with the two pure alignment kernels without sub-pixel local refinement. For the resolution assessment, we use the half-maximal Fourier shell correlation (FSC) criterion (Böttcher et al., 1997; Malhotra et al., 1998; van Heel, 1987) between the 3D reconstruction and the 5.0 Å low-pass filtered X-ray structure.

As illustrated in Fig. 3A, at SNR of 0.04 the FRM map resolution is 8.0 Å, whereas the SCF map resolution is 10.6 Å, so FRM yields a significant 2.6 Å resolution improvement. At higher SNR, such as 0.06 and 0.15, the FRM kernel still outperforms SCF (2 Å resolution improvement). These considerable resolution improvements clearly demonstrate the better image alignment afforded by FRM.

3.2.2. Effect of local refinement

The FSC curves in Figs. 3D–F demonstrate that the additional sub-pixel local refinement improves the
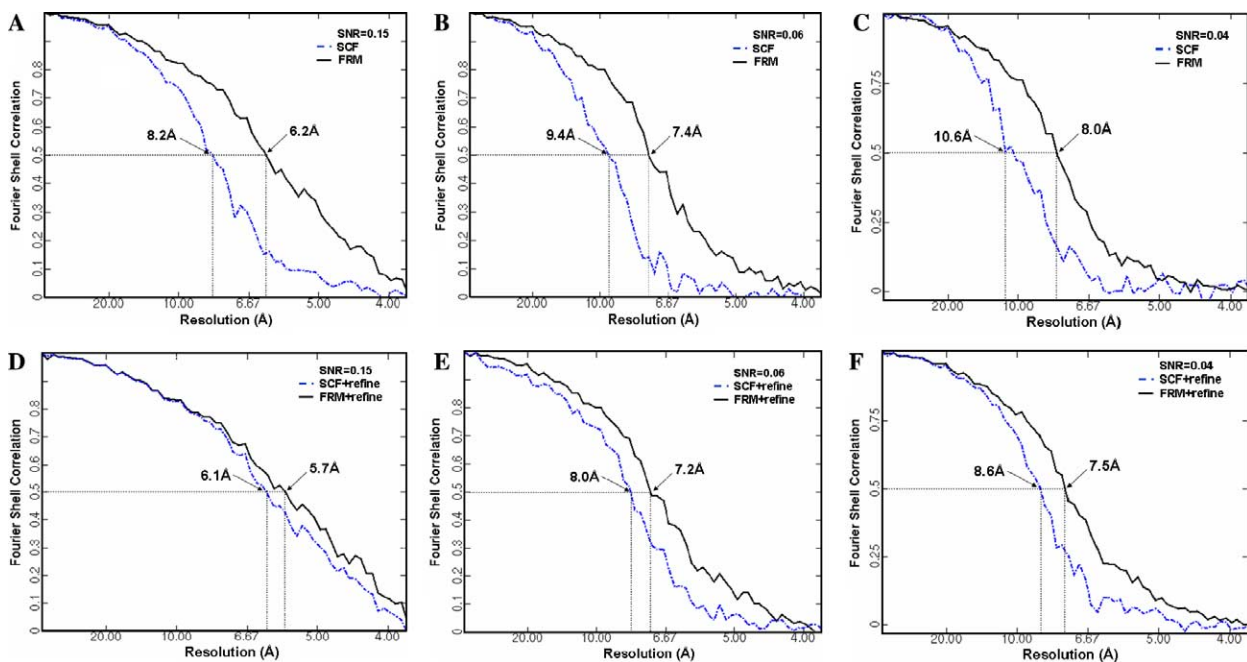

Fig. 3. GroEL 3D reconstructions at different noise levels utilizing FRM and SCF alignment kernels with and without additional sub-pixel refinement. (A–C) FSC between the (FRM or SCF) reconstruction and the low-pass filtered X-ray structure (see text) with SNR of 0.15, 0.06, and 0.04, respectively. (D–F) FSC between the (FRM + refine or SCF + refine) reconstruction and the low-pass filtered X-ray structure (see text) with SNR of 0.15, 0.06, and 0.04, respectively.

resolution for both kernels, especially for SCF. For instance, at SNR 0.06, the local refinement increases the FRM resolution from 7.4 to 7.2 Å, while for SCF, a considerable improvement from 9.4 to 8.0 Å was observed. Nevertheless, FRM + refine consistently yields higher resolutions than SCF + refine, e.g., at SNR 0.04, FRM + refine outperforms SCF + refine by 1.1 Å, whereas pure FRM originally outperformed SCF by 2.6 Å.

3.3. 3D reconstructions from experimental data

In this section, we use experimental electron micrographs of RDV to validate the realistic performance of FRM in the 3D reconstruction process. RDV is a double-shelled icosahedral particle with approximately 700 Å diameter. The 3D structure of RDV was determined at 6.8 Å with 3261 unique particles in the final reconstruction

Fig. 4. Representative RDV raw images (from close-to-focus micrographs) collected in a JEOL 4000 electron microscope. Image size is 300 × 300 pixels with 2.8 Å/pixel.

by single-particle electron microscopy (Zhou et al., 2001). Later on, an X-ray structure was reported (PDB code 1UF2, Nakagawa et al., 2003). The availability of both the EM 3D reconstruction and the X-ray structure makes RDV an appropriate model to validate the reconstruction performance of the alignment algorithm. Only part of the previous experimental image data from Zhou's reconstruction was used, also different reconstruction packages were employed in our reconstruction and the previous one (Zhou et al., 2001). It should be pointed out that this is the first 3D reconstruction of icosahedral particles using EMAN at sub-nanometer resolution (cf. Section 2).

3.3.1. Resolution estimation

In contrast to the use of focal pair images in the previous reconstruction (Zhou et al., 2001), we used only the close-to-focus micrographs in the following reconstruction tests. The close-to-focus images, as illustrated in Fig. 4, are extremely noisy but they do contain more high resolution information. Five rounds of iteration were carried out for both FRM and SCF kernels with 1.4° angular sampling precision. In addition, the sub-pixel local refinement function was turned on for both reconstructions. We estimate the final resolution by the half-maximal FSC between two models generated from half of the data set (Böttcher et al., 1997; Malhotra et al., 1998; van Heel, 1987).
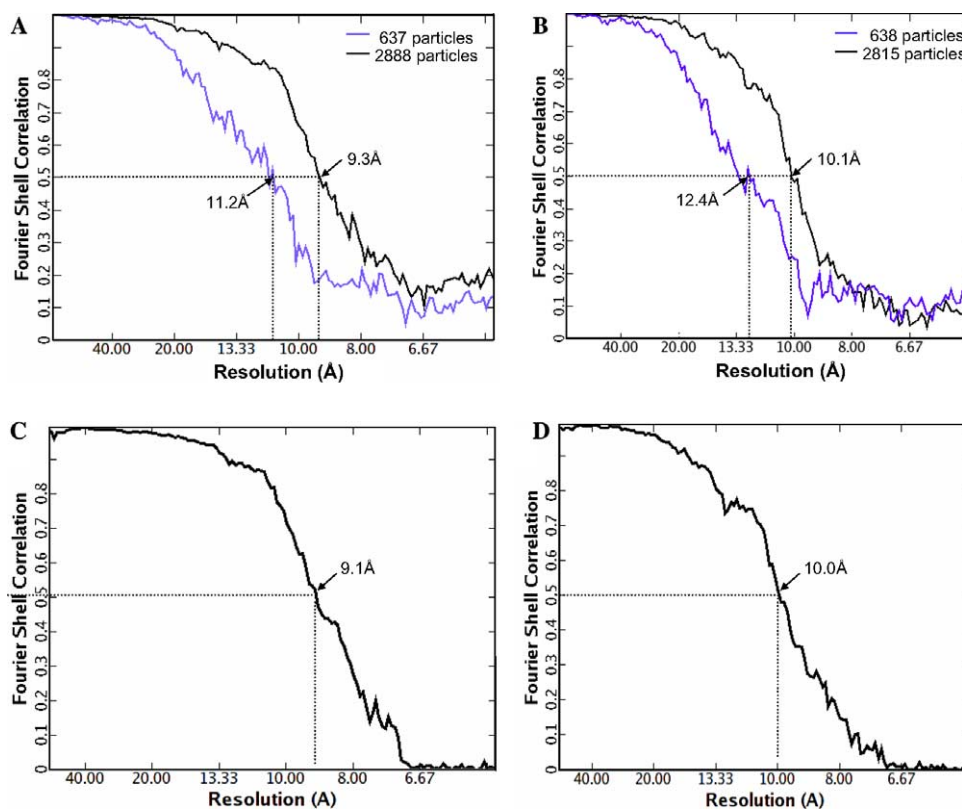

Fig. 5. RDV resolution assessment. (A and B) Two estimates of the final FRM + refine/SCF + refine reconstruction resolutions by *T* test. The effective resolution improved with increasing number of unique particles in the final reconstruction. The resolution estimate in (C and D) shows the FSC between the FRM + refine/SCF + refine reconstruction model and the previous model (Zhou et al., 2001), whose results are in accordance with the *T* test in (A and B). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this paper.)

Our test was conducted in two steps. In the first step, the same data set containing 1100 raw particles was used for both reconstructions. Six hundred and thirty-seven unique raw particles were used in the FRM + refine final reconstruction leading to a model with the effective resolution 11.2 Å (blue line in Fig. 5A); while for the SCF + refine reconstruction with 638 unique particles, the resolution was estimated at 12.4 Å (blue line in Fig. 5B), which is 1.2 Å lower than that of the FRM + refine map. In the second step, while using a larger data set containing 3500 particles, the effective resolution of FRM + refine

reconstruction increased to 9.3 Å with 2888 unique particles involved in the final reconstruction (black line in Fig. 5A); whereas for the SCF + refine map the resolution was estimated at 10.1 Å with 2815 unique particles from the same data set (black line in Fig. 5B). Both of the final structures from the two methods were masked using the iterative auto-masking procedure *automask2* in EMAN program*proc3d* with the mask extended by 6 voxels (16.8 Å) from the typical rendered isosurface (Ludtke et al., 2001, 2004). Furthermore, to validate these resolution assessments, we calculated the FSC between our
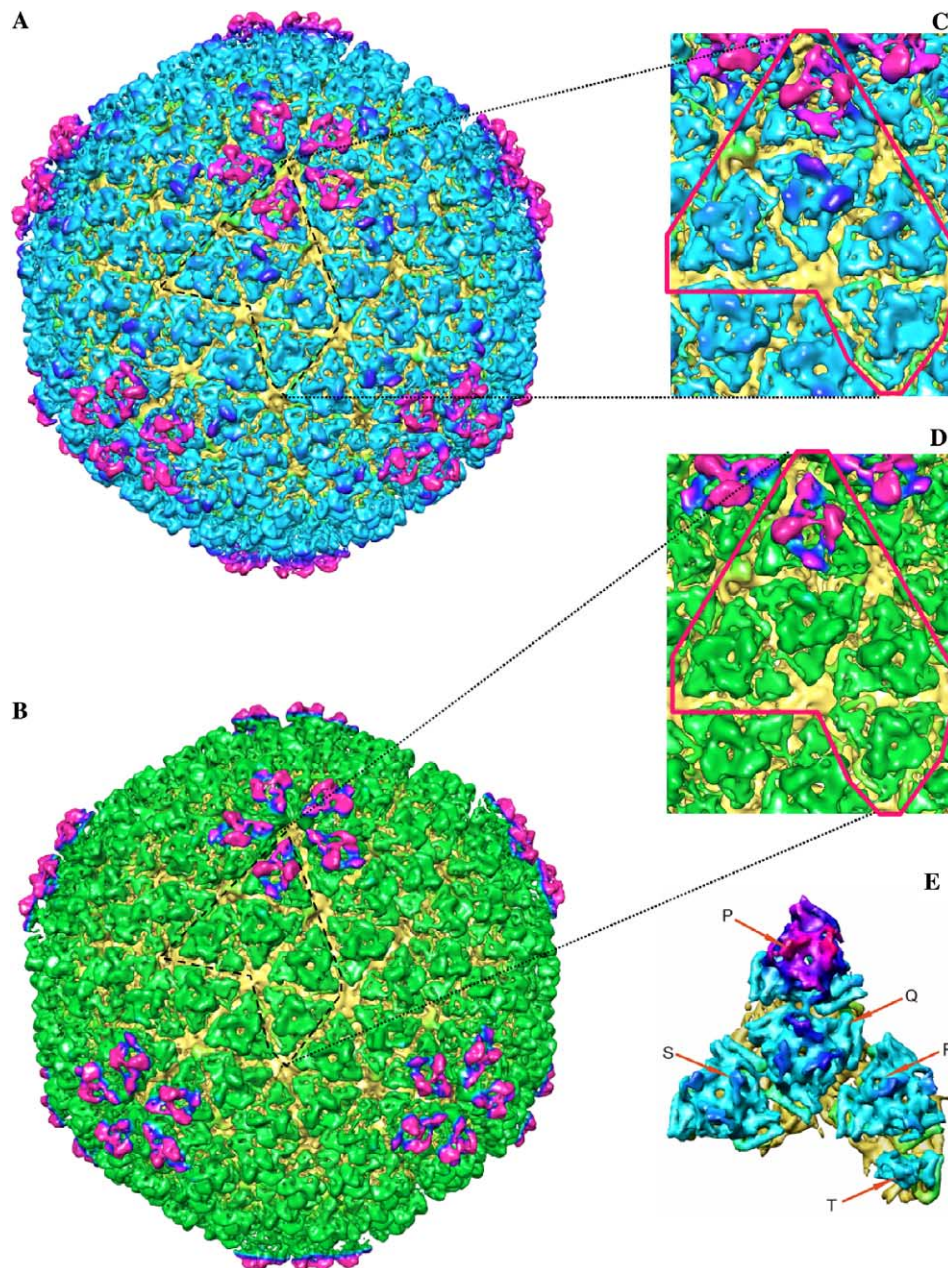

Fig. 6. Surface representations of the reconstructed RDV density maps. (A) FRM + refine reconstruction model contoured at 2.4σ. (B) SCF + refine reconstruction model contoured at 2.4σ. The regions within the purple frame in (C) and (D) are the asymmetric units enlarged from the FRM + refine and SCF + refine reconstructions shown in (A) and (B). (E) The asymmetric unit from the 8.4 Å resolution low-pass filtered X-ray structure, indicating the secondary structural features. Arrows point to the four 1/3 unique trimers, including one P, Q, R, S and 1/3 of T. All 3D structures in this paper were rendered by UCSF Chimera (Pettersen et al., 2004).

FRM + refine model and the previous model (Zhou et al., 2001). As illustrated in Figs. 5C and D, for FRM + refine the limiting FSC between the two maps is 9.1 Å (Fig. 5C), while for SCF + refine the limit is 10.0 Å (Fig. 5D). The resolution estimates show good agreement, and verify that in dealing with very noisy experimental EM data, FRM + refine still outperforms SCF + refine, yielding a slightly (about 1 Å) higher resolution in this example.

3.3.2. Reconstruction validation

To further validate the reliability of our reconstructions, we visually compared the FRM + refine and SCF + refine reconstructions (Figs. 6A and B) with the previous reconstruction (Fig. 1D in Zhou et al., 2001). In addition, we illustrate the outer shell P8 trimer subunits in the icosahedral asymmetric unit from our reconstructions (Figs. 6C and D) and from the X-ray structure (PDB entry 1UF2) which was low-pass filtered to a slightly higher 8.4 Å resolution to better illustrate the structural features (Fig. 6E). A visual comparison between the overall maps (Figs. 6A and B) as well as a detailed inspection of the asymmetric P8 trimer subunits (Figs. 6C–E) reveals that the overall shape and symmetry of FRM and SCF reconstructions are very similar to that of the previous reconstruction (Zhou et al., 2001), and both have very well revealed the outer shell icosahedral symmetry with the $T = 13l$ icosahedral lattice.

The fact that the outlines of helical regions can still be recognized down to the resolutions of $1/7$ to $1/10 \text{ Å}^{-1}$ (Chiu et al., 2005; Frank, 1996) permits us to quantify the FRM + refine reconstruction in an additional way. As illustrated in Fig. 7, by utilizing the Situs package (Chacón and Wriggers, 2002) we docked the P8 monomer X-ray structure (segmented from PDB entry 1UF2) into the corresponding P8 monomer density map segmented from the full FRM + refine reconstruction map (Jiang et al., 2001; Ludtke et al., 1999). From the superimposed map, especially the regions within the orange circles illustrated in Fig. 7, one can clearly see that the three turn α-helix (Fig. 7A) and the connected five and three turn α-helixes (Fig. 7B) were all very well fitted into the sausage-shaped densities.

The above visual inspection and docking test clearly demonstrate that utilizing the FRM alignment kernel can accomplish reliable 3D reconstructions directly from extremely noisy experimental data sets even in the absence of focal pairs.

4. Conclusion

In the current work, we fully integrated our recently developed FRM method as an alignment function into the single-particle analysis package EMAN. We have revisited the alignment of images with FRM to focus for the first time on the 3D reconstruction performance.

In reference-based reconstruction, if a raw image cannot be assigned to the correct reference class or a near neighbor class, it will either be discarded later in the class averaging procedure, or worse, pollute the final reconstruction. Consequently, the accuracy of the classification procedure, which heavily depends on the performance of the alignment algorithm employed, is one of the key factors to ensure a successful 3D reconstruction. Therefore, we first tested the classification accuracy on simulated GroEL data. Our examinations demonstrate that FRM exhibits 13% better classification accuracy with much narrow classification error distribution range than SCF (21 vs. 67) at a relative low SNR of 0.04. Unlike the SCF method which ignores the important phase information during calculating the rotational angle, the FRM method retains both amplitude and phase information, avoiding the intrinsic sensitivity to noise associated with the SCF method.
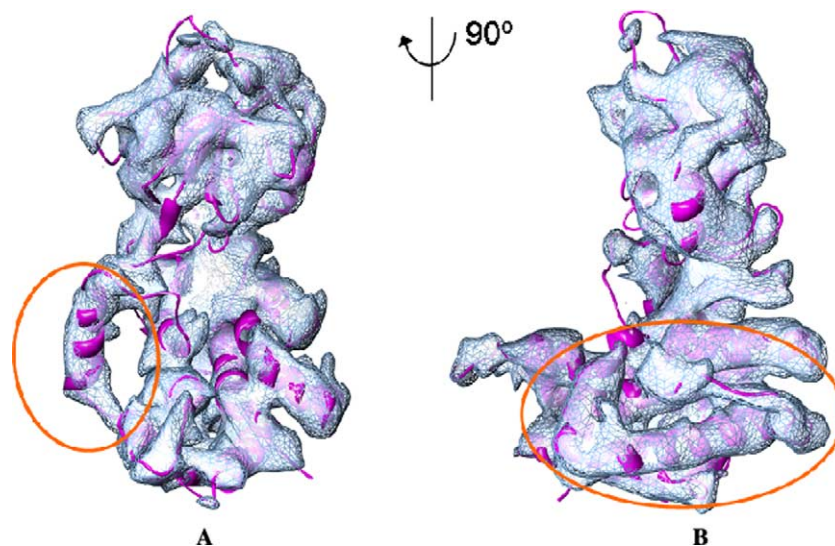

Fig. 7. Outer shell protein P8 extracted from the FRM + refine reconstruction (shown in Fig. 6A) compared to a P8 monomer atomic structure (PDB code 1UF2). (A) P8 monomer density map (contoured at 0.9σ) and a ribbon diagram of the X-ray structure. (B) Panel (A) rotated 90° in the clockwise direction. In (A) and (B), the orange circles highlight the well fitted α-helix positions.

The GroEL 3D reconstruction tests at various noise levels show that a considerable resolution improvement is achieved by FRM compared to SCF method in the absence of local sub-pixel refinement. For example, at an SNR of 0.04, FRM outperforms SCF by 2.6 Å with improvement from 10.6 to 8.0 Å resolution, which clearly validates the robustness of FRM at high noise. With local refinement the performance advantage is less pronounced but still significant. Accordingly, employing a sub-pixel local refinement after the initial alignment is a meaningful choice to improve alignment accuracy; however, one needs to bear in mind that a local refinement can improve the alignment accuracy only if the initial alignment was successful.

The reconstructions based on the experimental RDV EM images also support that FRM is a reliable and robust alignment kernel that provides a resolution gain in the 3D reconstruction procedure. The observed gain in resolution (9–11%) is significant. We have demonstrated that simply by utilizing a better algorithmic approach the resolution of 3D maps could be enhanced, in principle, even years after the data were collected. In summary, FRM is a powerful novel alignment kernel that will be useful in applications where the experimental images are extremely noisy.

Acknowledgments

References

Baker, T.S., Cheng, R.H., 1996. A model-based approach for determining orientations of biological macromolecules imaged by electron microscopy. J. Struct. Biol. 116, 120–130.

Baker, T.S., Olson, N.H., Fuller, S.D., 1999. Adding the third dimension to virus life cycles: three-dimensional reconstruction of icosahedral viruses from electron micrographs. Microbiol. Mol. Biol. Rev. 63, 862–922.

Baldwin, P.R., Penczek, P., Pawel, A., 2005. Estimating alignment errors in sets of 2-D images. J. Struct. Biol. 150, 211–225.

Böttcher, B., Wynne, S.A., Crowther, R.A., 1997. Determination of the fold of the core protein of hepatitis B virus by electron microscopy. Nature 386, 88–91.

Chacón, P., Wriggers, W., 2002. Multi-resolution contour-based fitting of macromolecular structures. J. Mol. Biol. 317, 375–384.

Chiu, W., Baker, M.L., Jiang, W., Dougherty, M., Schmid, M.F., 2005. Electron microscopy of biological machines at subnanometer resolution. Structure 13, 363–372.

Cong, Y., Kovacs, J.A., Wriggers, W., 2003. 2D fast rotational matching for image processing of biophysical data. J. Struct. Biol. 144, 51–60.

Frank, J., 1996. In: Three-Dimensional Electron Microscopy of Macromolecular Assemblies. Academic Press, New York, p. 260.

Frank, J., Goldfarb, W., Eisenberg, D., Baker, T.S., 1978. Reconstruction of glutamine synthetase using computer averaging. Ultramicroscopy 3, 283–290.

Frank, J., Radermacher, M., Penczek, P., Zhu, J., Li, Y., Ladjadj, M., Leith, A., 1996. SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. J. Struct. Biol. 116, 190–199.

Jiang, W., Baker, M.L., Ludtke, S.J., Chiu, W., 2001. Bridging the information gap: computational tools for intermediate resolution structure interpretation. J. Mol. Biol. 308, 1033–1044.

Joyeux, L., Penczek, P.A., 2002. Efficiency of 2D alignment methods. Ultramicroscopy 92, 33–46.

Kovacs, J.A., Chacon, P., Cong, Y., Metwally, E., Wriggers, W., 2003. Fast Rotational Matching of rigid bodies by fast Fourier transform acceleration of five degrees of freedom. Acta Crystallogr. D 59, 1371–1376.

Ludtke, S.J., Baldwin, P.R., Chiu, W., 1999. EMAN: Semiautomated software for high-resolution single-particle reconstructions. J. Struct. Biol. 128, 82–97.

Ludtke, S.J., Chen, D.H., Song, J.L., Chuang, D.T., Chiu, W., 2004. Seeing GroEL at 6 Å resolution by single particle electron microscopy. Structure 12, 1129–1136.

Ludtke, S.J., Jakana, J., Song, J.L., Chuang, D.T., Chiu, W., 2001. A 11.5 Å single particle reconstruction of GroEL using EMAN. J. Mol. Biol. 314, 253–262.

Malhotra, A., Penczek, P., Agrawal, R.K., Gabashvili, I.S., Grassucci, R.A., Junemann, R., Burkhardt, N., Nierhaus, K.H., Frank, J., 1998. *Escherichia coli* 70 S ribosome at 15 Å resolution by -electron microscopy: localization of fMet-tRNA and fitting of L1 protein. J. Mol. Biol. 280, 103–116.

Nakagawa, A., Miyazaki, N., Taka, J., Naitow, H., Ogawa, A., Fujimoto, Z., Mizuno, H., Higashi, T., Watanabe, Y., Omura, T., Cheng, R.H., Tsukihara, T., 2003. The atomic structure of rice dwarf virus reveals the self-assembly mechanism of component proteins. Structure 11, 1227–1238.

Penczek, P., Radermacher, M., Frank, J., 1992. Three-dimensional reconstruction of single particles embedded in ice. Ultramicroscopy 40, 33–53.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., Ferrin, T.E., 2004. UCSF chimera—a visualization system for exploratory research and analysis. J. Comput. Chem. 25, 1605–1612. <http://www.cgl.ucsf.edu/chimera/>.

Schatz, M., van Heel, M., 1992. Invariant recognition of molecular projections in vitreous ice preparations. Ultramicroscopy 45, 15–22.

van Heel, M., Harauz, G., Orlova, E.V., Schmidt, R., Schatz, M., 1996. A new generation of the IMAGIC image processing system. J. Struct. Biol. 116, 17–24.

van Heel, M., Schatz, M., Orlova, E., 1992. Correlation functions revisited. Ultramicroscopy 46, 307–316.

van Heel, M., 1987. Similarity measures between images. Ultramicroscopy 21, 95–100.

Zhou, Z.H., Baker, M.L., Jiang, W., Dougherty, M., Jakana, J., Dong, G., Lu, G., Chiu, W., 2001. Electron microscopy and bioinformatics suggest protein fold models for rice dwarf virus. Nat. Struct. Biol. 8, 868–873.