Available online at www.sciencedirect.com

Journal of Structural Biology 152 (2005) 104–112

Journal of Structural Biology

www.elsevier.com/locate/yjsbi

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*

Received 24 March 2005; received in revised form 25 August 2005; accepted 26 August 2005
Available online 30 September 2005

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.

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).

1047-8477/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.jsb.2005.08.006

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 subdegree 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:

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

where $\overline{\hat{f}_m(r)}$ is the complex conjugate of the Fourier coefficient of object f, $(\hat{h}^n_{r,\rho})_m$ 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 $\sigma^2_{\mathrm{signal}}/\sigma^2_{\mathrm{noise}}$, 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.