Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Structural Biology 144 (2003) 51–60
Journal of Structural Biology
www.elsevier.com/locate/yjsbi

2D fast rotational matching for image processing of biophysical data

Yao Cong,[a] Julio A. Kovacs,[b] and 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] Department of Molecular Biology, The Scripps Research Institute, 10550 N. Torrey Pines Rd., La Jolla, CA 92037, USA

Received 10 June 2003, and in revised form 25 August 2003

Abstract

In 3D single particle reconstruction, which involves the translational and rotational matching of a large number of electron microscopy (EM) images, the algorithmic performance is largely dependent on the efficiency and accuracy of the underlying 2D image alignment kernel. We present a novel fast rotational matching kernel for 2D images (FRM2D) that significantly reduces the cost of this alignment. The alignment problem is formulated using one translational and two rotational degrees of freedom. This allows us to take advantage of fast Fourier transforms (FFTs) in rotational space to accelerate the search of the two angular parameters, while the remaining translational parameter is explored, within a limited range, by exhaustive search. Since there are no boundary effects in FFTs of cyclic angular variables, we avoid the expensive zero padding associated with Fourier transforms in linear space. To verify the robustness of our method, efficiency and accuracy tests were carried out over a range of noise levels in realistic simulations of EM images. Performance tests against two standard alignment methods, resampling to polar coordinates and self-correlation, demonstrate that FRM2D compares very favorably to the traditional methods. FRM2D exhibits a comparable or higher robustness against noise and a significant gain in efficiency that depends on the fineness of the angular sampling and linear search range.

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

1. Introduction

In recent years, progress in single particle analysis of electron microscopy (EM) images was fueled by ever more sophisticated and powerful techniques that resulted in a wealth of medium- to low-resolution reconstructions (Ruprecht and Nield, 2001). The 2D alignment that determines the relative orientations and positions of two images remains a fundamental performance bottleneck for the 3D reconstruction process (Joyeux and Penczek, 2002). The two essential subprocedures in single particle analysis, 2D classification and 3D refinement, both require the computationally expensive alignment of a set of particle images with certain reference images. The required number of particle images increases dramatically as a function of desired resolution. For example, to achieve a resolution above $1/4\,\text{Å}^{-1}$ in cryo-EM, the image number is supposed to reach 10^6 (Glaeser, 1999; Henderson, 1995). Consequently, the efficiency of the 2D alignment method becomes a limiting factor for the improvement of resolution. Besides, due to the iterative procedure, small errors in the 2D alignment may result in inaccurate 3D reconstructions. Therefore, the overall quality of the 3D reconstruction relies heavily on the accuracy of the 2D alignment kernel used.

Various 2D alignment kernels have been proposed for single particle analysis. A leading method is Resampling to Polar Coordinates (RPC), which projects both images into polar coordinate space with respect to selected origin locations within the image frame (Joyeux and Penczek, 2002; Penczek et al., 1992). In this strategy, the rotational angle is determined from a 1D fast Fourier transform (FFT), which takes advantage of the Fourier convolution theorem, whereas the two translational parameters are determined by a limited exhaustive search. Another efficient alternative is the indirect alignment using the self-correlation function (SCF), which takes advantage of the fact that the SCF of an

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

1047-8477/$ - see front matter © 2003 Elsevier Inc. All rights reserved.
doi:10.1016/j.jsb.2003.09.017

image is invariant with respect to the translation of this image (Frank et al., 1978; Schatz and van Heel, 1992; van Heel et al., 1992). Thus, the search for the rotational and translational parameters can be split into two parts. First, the rotation angle between SCFs of the pair of images is obtained in the same way as with the RPC method, but instead of using the standard cross correlation function (CCF) that is maximized in RPC, the mutual-correlation function (MCF) is used as a matching criterion (van Heel et al., 1992). Second, the translational parameters are obtained by a 2D FFT, which takes advantage of the Fourier convolution theorem. Besides, there are several other widely used 2D alignment methods, including direct alignment using 2D FFT, sinograms (Radermacher, 1994), and direct alignment in real space. Among these traditional methods, according to (Joyeux and Penczek, 2002), RPC is the most accurate one. As to the efficiency, if the translational search can be restricted to a predetermined small value, RPC is the most efficient one; otherwise, SCF method ranks first (Joyeux and Penczek, 2002).

All these traditional methods describe the relative alignment of the images by one rotational and two translational parameters. In this work we express the problem in a different way: we propose the use of only one translational but two rotational parameters. We present a novel 2D fast rotational matching method (FRM2D) that takes advantage of this formulation. It allows us to handle the two rotational degrees of freedom (DOF) using a 2D FFT, avoiding the expensive zero padding required for a numerically stable Fourier convolution in linear space (Press et al., 1992), thereby significantly speeding up the matching process. FRM2D is the 2D analogue to our 6D FRM method, which was developed for fast rigid-body docking in 3D space using six DOF (Kovacs et al., 2003).

To test the performance and robustness of FRM2D, we carried out efficiency and accuracy comparisons with the two major traditional methods, RPC and SCF, on synthesized RNA polymerase images (Darst et al., 2002) over a range of noise levels. The results demonstrate that in low signal to noise ratio (SNR), FRM2D is comparable to the most accurate among the traditional 2D alignment methods. Also, FRM2D has the potential to significantly outperform the traditional methods with a gain in efficiency depending on the fineness of angular sampling and linear search range.

2. Methods

2.1. Theory

We describe two objects in two-dimensional space $\mathbf{R}^2$ by density functions:

$f : \mathbf{R}^2 \to \mathbf{R}$ and $g : \mathbf{R}^2 \to \mathbf{R}$.

These functions are assumed to be bounded (meaning that there is a constant $M > 0$ such that $|f(p)| < M$ for all p in $\mathbf{R}^2$) and of "compact support," i.e., they vanish outside a bounded set. When resampled to polar coordinates by means of bilinear interpolation, the density functions become functions of the radius r (distance to the reference point of each object) and the polar angle β: $f(r,\beta)$ and $g(r,\beta)$. These functions can be expended in Fourier series:

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

$$g(r,\beta) = \sum_{m=1-B}^{B-1} \hat{g}_m(r) e^{im\beta}, \tag{2}$$

where for each r, $\hat{f}_m(r)$ is the Fourier coefficient of $f(r,\beta)$ as a function of β keeping r fixed. To simplify our notation we choose m here to be a discrete reciprocal space coordinate corresponding to the continuous real-space angle β. The 'bandwidth' B, i.e., the number of frequencies used in the truncated Fourier expansions, is related to the sampling for the angular parameters. According to the sampling theorem, the number of sampling points in each circle equals $2B$ (Healy et al., 1998). In Fig. 1 we illustrate the idea and notations of FRM2D.

In order to conveniently express the rotation of the object, we define the rotation operator Λ_ϕ by:

$(\Lambda_\phi g)(p) := g(\phi^{-1}(p))$ for all points p in $\mathbf{R}^2$.

This operator rotates the image by an angle ϕ about the origin. We also define, for any $\rho \geqslant 0$, the translation operator T_ρ:

$(T_\rho g)(x,y) := g(x-\rho, y)$ for all (x,y) in $\mathbf{R}^2$.

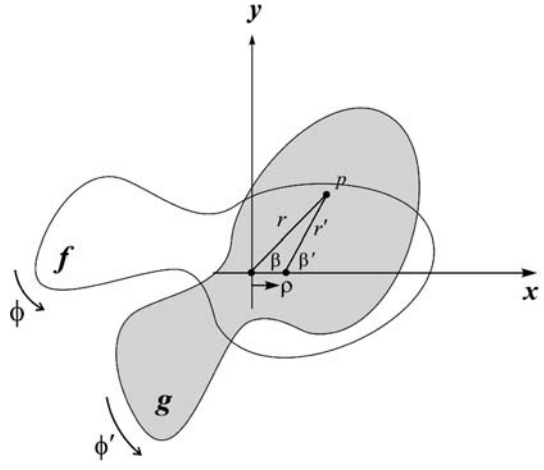

Fig. 1. FRM2D matching setup. r is the distance of a generic point p in the image to the reference point of object f (located at the origin of the coordinate system), and β is the corresponding polar angle. r' and β' are the analog quantities but relative to the reference point of object g located at $(\rho, 0)$. See text for details.

This operator shifts the image by a distance ρ along the positive x axis.

As shown in Fig. 1, in this algorithm we rotate both objects about their own reference point while translating one of them along the positive x axis only. The reference point could be any point within a given image, but it is convenient to choose the point to be an approximation of the corresponding center of mass (COM). The essential operations within the FRM2D scheme are as follows: firstly, translate one object (g in Fig. 1) along the positive x axis a distance ρ (Kovacs et al., 2003) while keeping the other one (f in Fig. 1) at the origin; secondly, rotate both of them around their own COMs (the COM of f is at the origin, while the COM of g is at $(\rho, 0)$) to find the best matching position at this value of ρ; then repeat the above two steps until reaching the best global score. Here we use two angular parameters ϕ and ϕ' to describe the rotations of the two objects f and g, respectively, about their own COMs. In this way, the three-dimensional search is performed over two angular parameters ϕ and ϕ' and one linear parameter ρ. Note that the following operations are carried out without any physical rotation or translation of the objects, so there are no interpolations involved other than those in the initial resampling to polar coordinates.

By using the expression of the two objects in Fourier space (Eqs. (1) and (2)) and the definition of the operators, the object f becomes, after rotation:

$$(\Lambda_\phi f)(r,\beta) = \sum_m \hat{f}_m(r) e^{im(\beta-\phi)}. \tag{3}$$

Likewise, the object g can be expressed, after translation and rotation, as:

$$(T_\rho \Lambda_{\phi'} g)(r,\beta) = (\Lambda_{\phi'} g)(r',\beta') = \sum_n \hat{g}_n(r') e^{in(\beta'-\phi')}, \tag{4}$$

where r' and β' are the analog quantities to r and β, but relative to the reference point of object g (located at $(\rho, 0)$, see Fig. 1).

The criterion for matching of the two objects is to maximize the correlation between one of them and a rotated and translated version of the other. In FRM2D, based on the above definition, the correlation between the rotated object f and the rotated and translated object g is written as:

$$c(\phi,\phi';\rho) = \int_{\mathbf{R}^2} \overline{\Lambda_\phi f} \cdot \overline{T_\rho \Lambda_{\phi'} g}. \tag{5}$$

Here, we perform the complex conjugation (denoted by overlines) of the two factors. This has no immediate effect since the factors are real, but we obtain later the desired sign in the exponential of the Fourier transform. The correlation is a function of the two rotations ϕ, ϕ' and of the distance ρ (Fig. 1). Since there is no boundary effect in FFTs in rotational space, we avoid the zero padding associated with the Fourier convolution in linear space. Using the expansions of the rotated and translated objects in Eqs. (3) and (4), the correlation in Eq. (5) turns out to be:

$$c(\phi,\phi';\rho) = \int_0^\infty \int_0^{2\pi} \sum_m \overline{\hat{f}_m(r)} e^{im(\phi-\beta)} \cdot \sum_n \overline{\hat{g}_n(r')} e^{in(\phi'-\beta')} r \, d\beta \, dr,$$

which can be rewritten as:

$$c(\phi,\phi';\rho) = \sum_{m,n} e^{i(m\phi+n\phi')} I_{mn}(\rho), \tag{6}$$

where

$$I_{mn}(\rho) = \int_0^\infty \left[\int_0^{2\pi} e^{-im\beta} \left(e^{-in\beta'} \overline{\hat{g}_n(r')} \right) d\beta \right] \cdot \overline{\hat{f}_m(r)} \cdot r \, dr.$$

If we denote the parenthesis inside the integral by $h^n_{r,\rho}(\beta) = e^{-in\beta'} \overline{\hat{g}_n(r')}$, we see that the inner integral is just the Fourier transform of $h^n_{r,\rho}(\beta)$ (in this equation we use the direct dependence of r' and β' on r, β, and ρ, see Fig. 1). Thus $I_{mn}(\rho)$ turns out to be:

$$I_{mn}(\rho) = 2\pi \int_0^\infty (\hat{h}^n_{r,\rho})_m \cdot \overline{\hat{f}_m(r)} \cdot r \, dr. \tag{7}$$

Eq. (6) shows that $I_{mn}(\rho)$ is just the 2D Fourier transform of the correlation function $c(\phi,\phi';\rho)$. Hence:

$$\hat{c}(m,n;\rho) = I_{mn}(\rho) = 2\pi \int_0^\infty (\hat{h}^n_{r,\rho})_m \cdot \overline{\hat{f}_m(r)} \cdot r \, dr. \tag{8}$$

Note that in Eq. (8), $(\hat{h}^n_{r,\rho})_m$ is the only factor containing information about object g, while the information about object f enters only through the second factor, $\hat{f}_m(r)$. Therefore, to match one reference image with many particle images, we compute $(\hat{h}^n_{r,\rho})_m$ for this reference image only once and store it in memory, and then use it for the matching with each of the particle images. This feature would in principle permit us to parallelize the code, although this was not exploited in the present work.

After computing $\hat{c}(m,n;\rho)$, an inverse 2D FFT gives us the cross-correlation function (CCF) as a function of the two rotation angles for each value of ρ. Then a peak search strategy is carried out to determine the maximum correlation value and the two corresponding rotation angles. We repeat this process for each ρ within a certain range. In practice, ρ will be small which allows for an efficient scan (see next section).

If the COMs of the pair of images can be determined exactly, we note that the two COMs should coincide, i.e., $\rho = 0$. In this case, the 2D Fourier transform of the correlation function $\hat{c}(m,n;\rho)$ in Eq. (8) reduces to the standard 1D Fourier transform:

$$\hat{c}_1(m) = 2\pi \int_0^\infty \hat{f}_m(r) \cdot \overline{\hat{g}_m(r)} \cdot r \, dr,$$

which is the way it is used in both RPC and SCF methods to determine the rotational angle.

2.2. Scanning of linear separation

The FRM2D strategy can be further accelerated if we restrict the scanning of the parameter ρ to a small range. In the following we show how to determine this range according to the shapes and sizes of the objects.

In this algorithm, g is considered as the reference image, and f as the particle image (Fig. 1). First we define the following radii: $r_{\min,f}$ and $r_{\max,f}$ are the min and max distances from the COM to the star hull of object f, respectively; similarly $r_{\min,g}$ and $r_{\max,g}$ are the corresponding min and max distances for object g.

In term of these, we define the maximum and minimum ρ (the distance between the COMs of the two objects) to be scanned:

$$\rho_{\max} = \begin{cases} r_{\max,g} - r_{\min,f} & (\text{if } r_{\max,g} > r_{\max,f}), \\ r_{\max,f} - r_{\min,g} & (\text{if } r_{\max,g} < r_{\max,f}), \end{cases}$$

$$\rho_{\min} = 0.$$

The expression for $\rho_{\max}$ indicates that if the two 2D images represent the same object (common case in EM image processing), and they have spherical shape, then the difference between $r_{\max,g}$ and $r_{\min,f}$ should be a small value. In this case, the range for ρ will be small, and the performance of the code will be faster than for generic particle shapes.

A comparison with the notation of the RPC method reveals that $\rho_{\max}$ is equivalent to the maximum translation of the image, k, except that k is measured in pixel units by (Joyeux and Penczek, 2002).

As mentioned earlier, another aspect that will improve the efficiency of FRM2D is the accuracy in the determination of the COMs of the images. Currently, in order to obtain the most accurate COMs of the low-resolution image, we use the averaged density of the reference images as a mask (Penczek et al., 1992). Moreover, in our implementation, if the maximum correlation values keep dropping for four consecutive ρ values, the scanning is aborted.

In addition, in Eq. (8) the integral range is restricted. Here are the actual maximum and minimum r values used in the computation of the integral:

$$r_{\max} = \begin{cases} \rho + r_{\max,g} & \text{if } (\rho + r_{\max,g}) < r_{\max,f}, \\ r_{\max,f} & \text{if } (\rho + r_{\max,g}) \geqslant r_{\max,f}, \end{cases}$$

$$r_{\min} = \begin{cases} \rho - r_{\max,g} & \text{if } (\rho - r_{\max,g}) \geqslant 0, \\ 0 & \text{if } (\rho - r_{\max,g}) < 0. \end{cases}$$

2.3. Elementary arithmetic operations

In this section we enumerate the elementary operations required by all three 2D alignment methods. This reveals how the image size, the angular sampling, and the image number affect the efficiency. In the 3D refinement process of single particle analysis, a small number of reference images are typically matched with a large number of particle images. Therefore, we only consider the number of operations of the "core" steps in each method, while ignoring the pre-computations which are determined just once. The arithmetic operation counts for the three 2D alignment methods are listed in Table 1.

We define the following elementary operations: rotate image (n^2), convert image to polar coordinates ($\sqrt{2}ns \cong 1.4ns$) (to make sure all the parts of the image are considered, the sampling radius is the diagonal of the image, $\sqrt{2}n$), 1D FFT ($n\log_2 n$ and $s\log_2 s$ for Cartesian and polar representations, respectively), and 2D FFT ($2n^2\log_2 n$ and $2s^2\log_2 s$ for Cartesian and polar representations, respectively). Here n is the size of the image in one dimension, and s is the number of angular sampling points in each concentric ring. To compare the effect of the angular sampling fineness for the three methods, in the following operation counts we

Table 1
The effect of the angular sampling and the exhaustive search range k on the operation counts of FRM2D, RPC, and SCF

k	s	FRM2D	RPC	SCF
		$(k+1)s^2(n/8+2\log_2 s)$	$lns(1.7+0.5\log_2 s)$	$n^2(32\log_2 n+38)+ns(0.5\log_2 s+1.7)$
4	2^5	117 760	1 132 190	2 744 090
	2^6	512 000	2 533 940	2 761 390
	2^7	2 211 840	5 607 010	2 799 330
	2^8	9 502 720	12 292 300	2 881 870
10	2^5	259 072	6 164 120	2 744 090
	2^6	1 126 400	13 795 900	2 761 390
	2^7	4 866 050	30 527 100	2 799 330
	2^8	20 906 000	66 924 700	2 881 870

$n = 104$ is the pixel size of the image in one dimension, s is the number of angular sampling points in each concentric ring, k is the exhaustive search range in pixels, and $l = (2k+1)^2$ is the number of possible translations. The second row lists the numbers of arithmetic operations for each method; rows 3–6 show the operation counts for each method with the COM uncertainty $k = 4$ pixels and under different angular samplings, namely 2^8, 2^7, 2^6, and 2^5 sampling points, corresponding to 1.4°, 3°, 6°, and 11° steps, respectively; while rows 7–10 show the operation counts with $k = 10$ pixels.

use the exact number of angular sampling points, s, instead of the approximation value πn used by Joyeux and Penczek (2002). As in (Joyeux and Penczek, 2002) we use k to express the maximum translation of the image in pixel units.

For the FRM2D counts, the one-time polar resampling and pre-computing of the 1D FFTs for both reference and particle images, and of $\hat{h}^n_{r,\rho}(m)$ for the reference images, are ignored. The computation of the integral in Eq. (8) requires $\frac{s}{2}\cdot\frac{n}{2}\cdot\frac{s}{2}=\frac{ns^2}{8}$ operations. Here one of the $s/2$ factors originates from the integral range of m. Since we use the FFTW package to calculate the Fourier transforms (FTs), where for the complex to real transform just the half dimension of the complex array is needed (Frigo and Johnson, 1997), only an $s/2$ range for m is considered in the implementation of our code. The $n/2$ factor originates from the integral range for r. The other $s/2$ factor reflects the integral range for n. Theoretically, n should vary from $-s/2$ to $s/2$, but since for the n-dependent term in the integral of Eq. (8), $\hat{h}^n_{r,\rho}(m)$, holds $\hat{h}^{-n}_{r,\rho}(m)=\hat{h}^n_{r,\rho}(m)=\hat{h}^n_{r,\rho}(-m)$, we can exploit this relationship and reduce the range of n to $s/2$. In the next step, an inverse 2D FFT is needed to calculate the CCF, which requires $2s^2\log_2 s$ operations. The above two steps must be repeated $(k+1)$ times for the translational scan of ρ, bringing the total operation count to $(k+1)s^2(n/8+2\log_2 s)$.

For the RPC counts, the one-time polar resampling of the reference images and the pre-computing of their 1D FFTs are ignored. The main computing effort is spent on the particle image resampling, $1.4ns$, followed by the 1D FFTs of all concentric rings of the particle images, $0.5n\cdot s\log_2 s$. These 1D FFTs must be multiplied by the 1D FFTs of the reference images, $ns/4$. Finally an inverse 1D FFT is carried out, $s\log_2 s$, but this term is small and is neglected in the following. The above processes require $(1.4ns+0.5ns\log_2 s+0.25ns)\cong ns(1.7+0.5\log_2 s)$ operations. These steps have to be repeated for $l=(2k+1)^2$ translations of the particle images, bringing the total number of operations to $lns(1.7+0.5\log_2 s)$. In the image processing of EM data, the value of $0.5\log_2 s$ is usually a single digit number, so we keep the constant 1.7 in this operation count, in contrast to (Joyeux and Penczek, 2002), where it is neglected.

For the SCF counts, we ignore the one-time precomputing of the SCFs for the reference and particle images. The angle between the two SCFs of reference and particle images is calculated using the resampling to polar coordinates technique as described above, requiring $ns(1.7+0.5\log_2 s)$ operations (no zero padding in this step). To calculate the translation parameters by using the convolution theorem, each image is padded with zeros to double the size in each dimension, therefore the operation number of 2D FFTs for the double sized image is $2(2n)^2\log_2 2n=8n^2\log_2 n+8n^2$. Then a rotation of the particle image by the angle found, and a 2D FFT of the image must be carried out, consuming $(n^2+8n^2\log_2 n+8n^2)$ operations. Subsequently, an integral and an inverse 2D FFT have to be performed to calculate the CCF, $(2n^2+8n^2\log_2 n+8n^2)$ operations. The above two steps must be performed for both the angle found and for that angle increased by 180°, using $n^2(32\log_2 n+38)$ operations. Therefore, the total number of operations required is $n^2(32\log_2 n+38)+ns(0.5\log_2 s+1.7)$.

3. Results

3.1. Test image preparation

Images observed by EM are not true projections of the specimen. Imaging artifacts include the effects of the contrast transfer function (CTF), which is introduced through electron lens aberrations and defocusing (Zhu et al., 1997; Ludtke and Chiu, 2002), and also the envelope function of the microscope, which contains contributions from a number of effects, such as spatial and temporal coherence, specimen motion, etc. (Hanszen, 1971). In addition, background noise is present from a variety of sources. Therefore, in the process of synthesizing test images, we attempted to closely emulate the image formation process in the electron microscope including the effects of CTF, envelope function and noise.

In this paper we take RNA polymerase as an example (Darst et al., 2002). All test images were synthesized from the atomic structure using the EMAN single particle analysis software (Ludtke et al., 1999). For the efficiency test (benchmark) of the three 2D alignment methods, we generated the top view 2D projection of RNA polymerase as a test image (see Fig. 2), where a Gaussian low-pass filter with half-width $1/15\,\text{Å}^{-1}$ was applied to simulate the effect of the envelope function. The image size is 104×104 pixels and the pixel size in this test is 1.7 Å/pixel. In addition, to make the test realistic, CTF and Gaussian noise were also applied to the image. Here, we use the CTF and noise parameters taken from the experimental measurements made on a JEOL 4000 electron microscope with a LB6 gun operated at 400 kV with an objective lens spherical aberration coefficient of 4.1 mm (Ludtke et al., 1999). Fig. 2B shows the representative projection with CTF parameters applied, which was used as reference image in the efficiency test. We rotated this image 100° and then translated it 6 and −4 pixels in the x and y directions, respectively, to provide a target for our search. To simulate the effect of background noise, Gaussian noise was applied to this rotated and translated image using EMAN's command "applyctf" with noise level 0.3. The resulting image, Fig. 2C, was used as the particle image

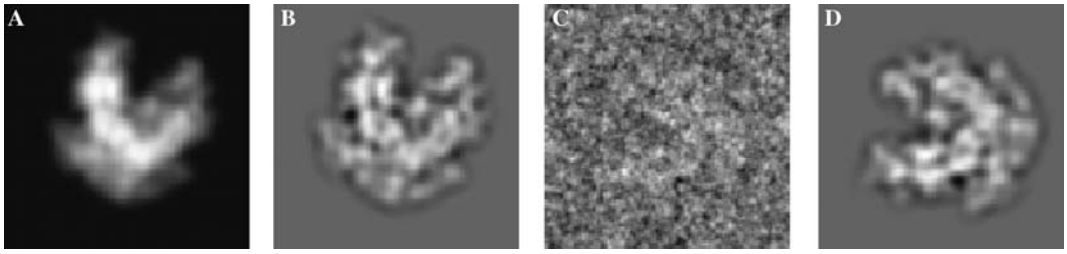

Fig. 2. Synthetic images used in the test of algorithm efficiency. (A) Projection of RNA polymerase (Darst et al., 2002) after Gaussian low-pass filtering with half-width of spatial frequency $1/15\,\text{Å}^{-1}$. (B) Image (A) after application of the CTF (see text). (C) Image (B) after rotation by 100° and translation by (6 pixels, −4 pixels), followed by corruption with Gaussian noise (SNR = 0.1477). (D) Image (B) but rotated and translated to be in the same location as image (C). We use image (D) as the reference image to calculate the SNR in (C).

in the benchmark calculations. We use Fig. 2D, which is similar to Fig. 2B, except that the object is in the location of Fig. 2C, as the reference to calculate the signal to noise ratio (defined as $\text{SNR} = \sigma^2_{\text{signal}}/\sigma^2_{\text{noise}}$). For example, the SNR value of Fig. 2C is 0.1477.

To carry out the accuracy comparison of the three 2D alignment methods, we generated six random 2D projections of RNA polymerase of various shapes (Fig. 3). A Gaussian low-pass filter with half-width of $1/15\,\text{Å}^{-1}$ was applied to all these projections to simulate the effect of the envelope function. Then a CTF simulation was applied to them. These images (after the application of envelope function and CTF) were used as reference images in the accuracy test (Fig. 3A–F). Subsequently, each reference image was rotated and translated with five different sets of rotation and translation values: (100°; 6 pixels, −4 pixels), (320°; 3 pixels, −7 pixels), (88°; −2 pixels, 3 pixels), (160°; 5 pixels, 3 pixels) and (76°; −3 pixels, −4 pixels). Finally, 100 different levels of Gaussian noise (from 0.01 to 1 in EMAN command "applyctf") were added to each of these rotated and shifted images. To provide reasonable statistics, we repeated the addition of (statistically independent) Gaussian noise 20 times for each image. Thus, a total of 60,000 images were generated as test images to compare the accuracy of the 2D alignment methods. We have also tested the method with the circularly shaped GroEL, obtaining correct alignment results. For the

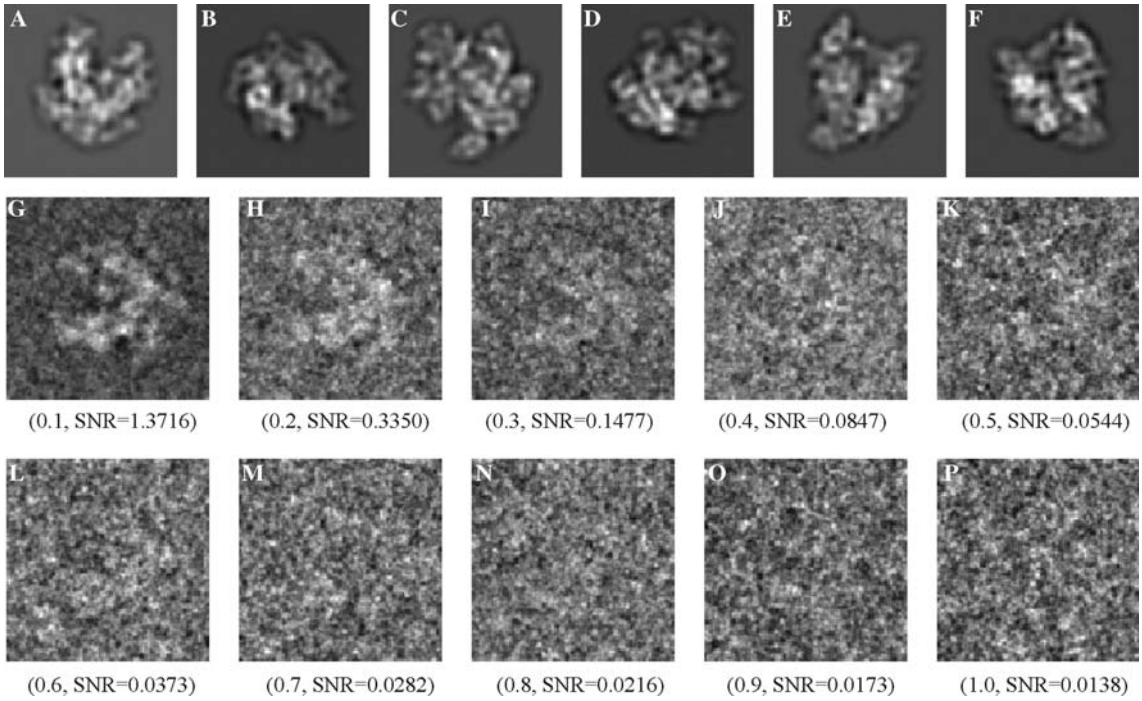

Fig. 3. Synthetic images used for the accuracy test. (A–F) Six random projections of RNA polymerase after Gaussian low-pass filtering with half-width $1/15\,\text{Å}^{-1}$, to which a CTF was applied, the resulting images being used as reference images. (G–P) Taking projection (A) as an example, image is rotated and translated (here: 100°, 6 pixels, −4 pixels), and then corrupted with Gaussian noise (noise level range from 0.01 to 1 using EMAN command "applyctf," with a total of 100 noise levels), leading to a variation in SNR from 0.008 to 153.12. Here we show ten examples with noise levels changing from 0.1 to 1 (in EMAN units), in 0.1 unit increments.

sake of brevity, here we just list the results for RNA polymerase.

3.2. Efficiency comparison

Among the traditional 2D alignment methods, if the translational exhaustive search range k in the RPC method can be restricted to a small value, RPC is faster than the other methods including SCF; otherwise, SCF ranks as the most efficient one (Joyeux and Penczek, 2002). To test the overall efficiency performance of FRM2D, we therefore implemented the fastest two methods, RPC and SCF, for comparison. In the test, the exhaustive search range k was restricted to 4 pixels for both FRM2D and RPC methods. If the COMs of the objects can be reasonably well determined, a search restricted to 4 pixels is sufficient and will not degrade the performance. Besides, for SCF, a zero padding to double the size in each image dimension was applied because of the requirement for the numerically accurate convolution in linear space (van Heel et al., 1992). All the computations were carried out on an AMD Athlon 1900+ dual processor PC under Linux.

To test the effect of the particle image number on the efficiency of the three methods, we used particle image numbers increasing from 10, 50, 100, 1000, 2000, 3000 to 4000, while the reference image number was restricted to 2 (to test the implementation of the loop for the reference images in the source code). Besides, to test the effect of the angular sampling on the efficiency, the tests were conducted with four angular sampling levels: 11°, 6°, 3°, and 1.4°. The timings for these angular samplings are given in Table 2. Due to the near-linear scaling we show the timings only for 10 and 4000 particle images.

As seen in Table 2, compared to RPC and SCF, FRM2D shows a more pronounced dependence on the fineness of the angular sampling. For example, in the case of 4000 particle images the increase in angular sampling from 6° to 3° yields an increase by a factor 3.1 for FRM2D, whereas the timings increase by a factor 2.1 and 1.1 for RPC and SCF, respectively. The operation numbers in Table 1 also indicate that for FRM2D, the efficiency depends quadratically on the sampling point numbers, s, and for RPC, the dependence is linear, while for SCF, the effect of s can be ignored.

The timings in Table 2 illustrate that for 4000 particle images ($k = 4$ pixels, $n = 104$ pixels) FRM2D is faster than RPC. This is in agreement with the corresponding comparison of the operation numbers listed in Table 1. An increasing search range k will attenuate the speed of RPC (quadratically) more than that of FRM2D (linearly). Likewise, an increasing image size affects more the performance of RPC. In contrast, an increased angular sampling slows FRM2D more than RPC. The combination of these factors determines which method is a better choice in different conditions. For example, for $k = 4$ and $n = 300$, FRM2D is the better choice for any reasonable angular sampling as fine as 0.7°.

As seen in Table 2, the comparison of FRM2D with SCF shows the importance of angular sampling. For instance, for 4000 particle images, in the coarse angular samplings of 11° and 6°, FRM2D is faster than SCF; while for the fine angular samplings of 3° and 1.4°, SCF is faster. This is in general agreement with the corresponding comparison of the operation numbers listed in Table 1. Unlike in SCF, there is no zero padding in FRM2D, which affords FRM2D more efficiency compared to SCF in coarse angular sampling. However, FRM2D determines the relative orientation of the objects through a 2D FFT technique, while SCF through a 1D FFT. In fine sampling this slows down FRM2D compared to SCF.

Next, we demonstrate the effect of the image number, indicating the effect of precomputation, on the efficiency of FRM2D and SCF alignment methods. Taking an angular sampling of 6° as an example, for which both methods show comparable efficiency (Table 2), Fig. 4 shows the compute time as a function of the number of image matches (reference image number times particle image number). The number of reference images used was 2 and 5. For SCF, the effect of precomputation is negligible, i.e., the lines keep the same linear trend irrespective of the reference image number. But in the case of FRM2D, the effect of the precomputation to determine $(\hat{h}^n_{r,\rho})_m$ for each reference image (Eq. (8)) is quite obvious: FRM2D (6°) lags behind SCF for low matching numbers, but it improves and outperforms SCF when the matching image number increases. Therefore, FRM2D is especially suitable for many particle images, since the extra

Table 2
The effect of particle image numbers ($n = 104$ pixles) and of the angular sampling on the efficiency of FRM2D, RPC, and SCF with a $k = 4$ pixel search range (times in seconds)

Angular sampling	10 Particle images			4000 Particle images		
	FRM2D	RPC	SCF	FRM2D	RPC	SCF
11°	0.31	1.15	0.51	79.98	447.02	193.39
6°	0.88	1.86	0.53	174.89	736.52	201.19
3°	3.15	3.96	0.57	542.93	1552.72	214.87
1.4°	12.34	7.38	0.64	2044.64	2875.23	238.34

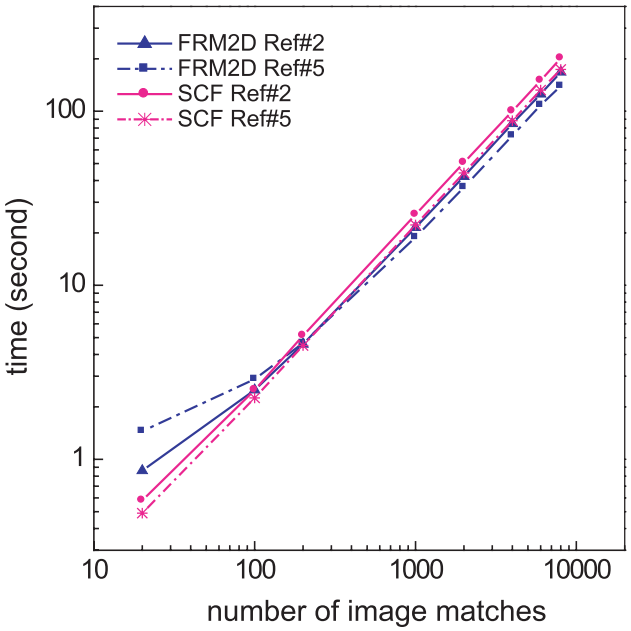

Fig. 4. 2D alignment time versus the number of image matches (reference image number times particle image number) for FRM2D and SCF methods. The reference image numbers are 2 and 5, and the angular sampling is 6°.

overhead to determine $(\hat{h}^n_{r,\rho})_m$ for each reference image becomes negligible.

3.3. Accuracy comparison

In this section, we describe our evaluation of the overall matching accuracy of the three 2D alignment methods. The pixel error $d\left|\sin\frac{\Delta\phi}{2}\right| + \sqrt{\Delta x^2 + \Delta y^2}$ was used to measure the accuracy (Joyeux and Penczek, 2002). Here d is the diameter of the particle in pixel units. In this formula, the first term describes errors introduced by the rotation misalignment $\Delta\phi$, and the second one describes the errors from the translation misalignment Δx and Δy. Fig. 3 shows a representative set of the test images. There were 100 Gaussian noise levels added by the EMAN software package, corresponding to an SNR range between 0.008 and 153.12. The total number of test images was 60,000. All the accuracy tests were carried out at an angular sampling level of 1.4°, fine enough to give highly accurate results.

The scatter plots of observed pixel errors versus SNR for FRM2D, RPC, and SCF are shown in Fig. 5A–C, respectively, and the average pixel error versus SNR for the three methods is shown in Fig. 5D. Fig. 5A–C show that for a high SNR range (>10) the pixel errors for all the three methods remains at a low value near 1 pixel, so in Fig. 5D we focus on the interesting range where the methods break down (SNR between 0.008 and 10).

As seen in Fig. 5D, for all the three methods, in the high SNR range (>0.2 for FRM2D and RPC, and >0.4 for SCF), the average pixel errors remain below 2 pixels; when the SNR decreases below this value (0.2 for FRM2D and RPC, and 0.4 for SCF), the average pixel errors increase sharply. In the case of the SCF method, the increase of pixel error with the decrease of SNR below 0.4 is more pronounced than that of FRM2D and RPC, suggesting that SCF exhibits more intrinsic sensitivity to noise than the other two alignment methods. In addition, we see in Fig. 5A, B, and D that for FRM2D and RPC, the overall shapes of their curves are very similar especially in the low SNR region, suggesting that the accuracies of FRM2D and RPC are quite comparable.

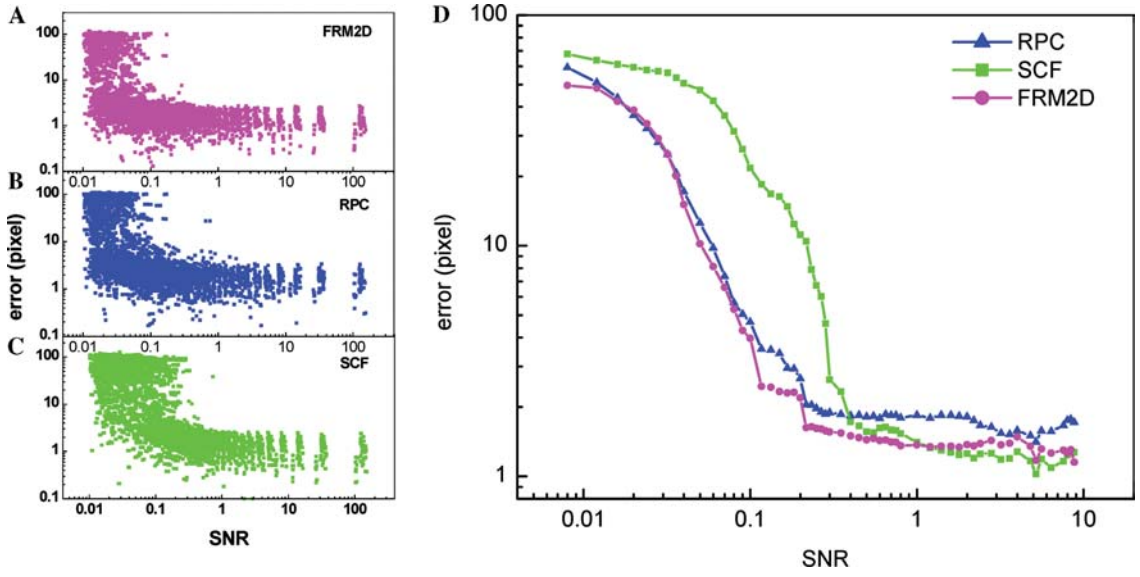

Fig. 5. (A)–(C) Scatter plot of pixel error as a function of SNR for the 2D alignment methods including FRM2D, RPC, and SCF. (D) Average pixel error as a function of SNR for the three methods.

To obtain accuracies better than 5 pixels, the minimum SNR values required are: 0.08, 0.09, and 0.28 for FRM2D, RPC, and SCF, respectively. This means that to obtain reliable alignment result, SCF usually requires much higher quality images than the other two methods. In the low SNR value range (<0.04) the alignment results are essentially random.

Besides the intrinsic sensitivity of SCF to noise, another factor that reduces SCF accuracy is interpolation after rotational alignment. Both translation by a non-integer number and any rotation (except for special cases of 90°, 180°, and 270°) require interpolation. In contrast, FRM2D and RPC do not suffer from the rotational interpolation problem, since the translations and rotations are performed analytically.

4. Conclusions

In this paper we propose a novel 2D fast rotational matching (FRM2D) algorithm to conduct 2D image matching in single particle analysis. By recasting the matching problem (usually including one rotational and two translational DOF) into a formulation involving two rotational and one translational DOFs, we are able to use FFT techniques to accelerate the two angular parameter searches, while the translational parameter is searched within a limited range.

To verify the robustness of FRM2D method, we performed efficiency and accuracy tests with synthesized RNA polymerase images over a range of noise levels. For comparison purposes, we implemented two other leading 2D alignment methods, RPC and SCF, which ranked best among all traditional alignment algorithms in overall efficiency and accuracy (Joyeux and Penczek, 2002).

We tested the accuracy of the three methods on a set of test images prepared according to the theory of image formation in the electron microscope. The overall accuracy of FRM2D is similar to the best performing traditional method, RPC, due to the use of the standard correlation and due to the absence of rotational resampling artifacts inherent to both methods. With decreasing SNR value, the accuracy of SCF deteriorates sooner than observed for the other two methods, which is in line with earlier observations (Joyeux and Penczek, 2002).

The efficiency comparison shows that FRM2D has the potential to outperform the traditional 2D alignment methods depending on the desired fineness of angular sampling and the exhaustive search range. The observed possible gains in efficiency will require future verification under realistic settings within established single particle reconstruction packages. Suffice it to note that FRM2D has a number of desirable performance properties: the translational search scales linearly, not quadratically, with the "particle picking inaccuracy," and the 2D FFT operates in rotational space, avoiding the costly zero-padding of linear space FFTs. Based on our efficiency and accuracy tests, FRM2D should be a sensible choice for image alignment, as it combines the speed of SCF with the accuracy of RPC.

FRM2D awaits further test when used with real EM images. At present, the FRM2D implementation is written in C, based in part on libraries of our Situs fitting package. Since Situs is limited to 3D modeling, FRM2D will be disseminated separately and the source code is currently available from us by request.

Acknowledgments

We thank Stefan Birmanns and Pablo Chacón for assistance with the implementation of the codes, Pawel A. Penczek and Joachim Frank for helpful criticism, and Steven J. Ludtke and Wen Jiang for the introduction to EMAN. Our work is supported by the La Jolla Interfaces in Science Program (J.K.) and by grants to W.W. from National Institutes of Health (1R01GM62968), Human Frontier Science Program (RGP0026/2003), and Alfred P. Sloan Foundation (BR-4297).

References

Darst, S.A., Opalka, N., Chacon, P., Polyakov, A., Richter, C., Zhang, G., Wriggers, W., 2002. Conformational flexibility of bacterial RNA polymerase. PNAS 99, 4296–4301.

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

Frigo, M., Johnson, S.G., 1997. The fastest Fourier transform in the West, for more information regarding FFTW and the software downloading, see http://www.fftw.org/.

Glaeser, R.M., 1999. Review: electron crystallography: present excitement, a nod to the past, anticipating the future. J. Struct. Biol. 128, 3–14.

Hanszen, K.J., 1971. The optical transfer theory of the electron microscope: fundamental principles and applications. In: Barer, R., Cosslett, V.E. (Eds.), Advances in Optical and Electron Microscopy, vol. 4. Academic Press, New York, pp. 1–84.

Healy Jr., D., Rockmore, D., Kostelec, P., Moore, S., 1998. FFTs for the 2-sphere-improvements and variations. Preprint and software available at www.cs.dartmouth.edu/~geelong/sphere/.

Henderson, R., 1995. The potential and limitations of neutrons, electrons and X-rays for atomic resolution microscopy of unstained biological molecules. Q. Rev. Biophys. 28, 171–193.

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. Cryst. D 59, 1371–1376.

Ludtke, S.J., Chiu, W., 2002. Image restoration in sets of noisy electron micrographs. In: Proceedings of the 2002 IEEE International Symposium on Biomedical Imaging. IEEE Cat 02EX608C, pp. 745–748.

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

Penczek, P., Radermacher, M., Frank, J., 1992. 3-Dimensional reconstruction of single particles embedded in ice. Ultramicroscopy 40, 33–53.

Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., 1992. Numerical recipes in C. In: The Art of Scientific Computing, second ed. Cambridge University Press, New York. Chapter 13.

Radermacher, M., 1994. 3-Dimensional reconstruction from random projections: orientational alignment via radon transforms. Ultramicroscopy 53, 121–136.

Ruprecht, J., Nield, J., 2001. Determining the structure of biological macromolecules by transmission electron microscopy, single particle analysis and 3D reconstruction. Prog. Biophys. Mol. Biol. 75, 121–164.

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

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

Zhu, J., Penczek, P.A., Schröder, R., Frank, J., 1997. Three-dimensional reconstruction with contrast transfer function correction from energy-filtered cryoelectron micrographs: procedure and application to the 70s *Escherichia coli* ribosome. J. Struct. Biol. 118, 197–219.