Available online at www.sciencedirect.com
ScienceDirect
Journal of Structural Biology xxx (2007) xxx–xxx
Journal of Structural Biology
www.elsevier.com/locate/yjsbi

Markov random field based automatic image alignment for electron tomography

Fernando Amat [a,1], Farshid Moussavi [a,1], Luis R. Comolli [b], Gal Elidan [c], Kenneth H. Downing [b], Mark Horowitz [a,*]

[a] Department of Electrical Engineering, Stanford University, Stanford, CA, USA
[b] Life Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA
[c] Department of Computer Science, Stanford University, Stanford, CA, USA

Received 14 March 2007; received in revised form 18 June 2007; accepted 10 July 2007

Abstract

We present a method for automatic full-precision alignment of the images in a tomographic tilt series. Full-precision automatic alignment of cryo electron microscopy images has remained a difficult challenge to date, due to the limited electron dose and low image contrast. These facts lead to poor signal to noise ratio (SNR) in the images, which causes automatic feature trackers to generate errors, even with high contrast gold particles as fiducial features. To enable fully automatic alignment for full-precision reconstructions, we frame the problem probabilistically as finding the most likely particle tracks given a set of noisy images, using contextual information to make the solution more robust to the noise in each image. To solve this maximum likelihood problem, we use *Markov Random Fields* (MRF) to establish the correspondence of features in alignment and robust optimization for projection model estimation. The resulting algorithm, called Robust Alignment and Projection Estimation for Tomographic Reconstruction, or RAPTOR, has not needed any manual intervention for the difficult datasets we have tried, and has provided sub-pixel alignment that is as good as the manual approach by an expert user. We are able to automatically map complete and partial marker trajectories and thus obtain highly accurate image alignment. Our method has been applied to challenging cryo electron tomographic datasets with low SNR from intact bacterial cells, as well as several plastic section and X-ray datasets.

Keywords: Cryo electron microscopy; Correspondence; Probabilistic inference; Tomography; Markov Random Fields; Alignment; Projection model estimation

1. Introduction

Cryo electron tomography (cryo-ET) is the primary method for obtaining 3D reconstructions of intact bacteria, viruses, and complex molecular machines. While electron tomography of plastic sections of eukaryotic cells and tissues is an established technique that has been used with great success in the 3D reconstruction of sub-cellular organelles and macromolecular assemblies, cryo-ET of whole cells is a relatively new field (Koster and Klumperman, 2003; Baumeister, 2002; Medalia et al., 2002; Koster et al., 1997; Jensen and Briegel, 2007). Two aspects of the study of whole cells by cryo-ET make it more challenging than working with plastic sections. First, biological material embedded in amorphous ice can tolerate much lower doses than the same material embedded in plastic. Second, the diameter of most bacteria of interest is considerably larger than the thickness of typical plastic sections. As a result the SNR in cryo-ET is much lower than in ET of plastic-embedded samples, and this directly leads to significant difficulty of precisely aligning the projections, with

* Corresponding author. Address: Computer Systems Laboratory, Room 306, Gates Computer Science Building, 353 Serra Mall, Stanford, CA 94305, USA. Fax: +1 650 725 6949.
[1] These authors contributed equally to this work.

1047-8477/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.jsb.2007.07.007

the potential for reducing the resolution of the resulting 3D reconstruction.

Before reconstruction, a projection model which defines the transformation from the 3D object space to the 2D images must be estimated. This model corrects for mechanical imperfections and distortions during data acquisition, and solves for the alignment of the images. High accuracy image alignment through the tracking of fiducial markers, generally small gold beads, and projection model estimation are key determinants for the quality of full-precision reconstructions.

Currently, the step that fails first in challenging datasets is finding corresponding markers through a series of images. This failure leads to incorrect trajectories which in turn throw off the projection model estimation, making significant manual intervention inevitable. For a full-precision reconstruction, it is not unusual for an expert user to spend up to a full day of manual feature identification and refinement involving several iterations on a dataset, sometimes not knowing until the very end whether a reconstruction was biologically relevant.

This requirement for extensive manual intervention is not only costly in terms of an expert user's time, but also is a bottleneck in high throughput generation of tomographic reconstructions. With better microscope stages and new software for automatic data acquisition, data is now more easily acquired than processed and analyzed, and this trend is only going to accelerate in the future. Full-precision automatic alignment software is needed to enable tomographic reconstructions to reach the pace of acquisition of automated datasets.

Extensive work has been published for automatic marker detection and correspondence (e.g. Kremer et al., 1996; Mastronarde, 1997; Ress et al., 1999; Brandt et al., 2001; Heymann, 2001; Brandt and Ziese, 2006; Frank, 2006; Zheng et al., 2007) which has performed well for many datasets. However, correspondence for low SNR cryo-ET images is still prone to error. For the most part, these methods attempt to correspond individual markers one at a time in adjacent images using assumptions about the underlying projection geometry. This approach works well when markers are well defined and sparsely located. However, if markers are close together and inaccurately localized, higher order projective distortions and other uncertainties may cause the program to choose the wrong correspondence for a given marker. Such a mistake might affect the estimated projection model, which in turn can affect the location of new trajectories. The result is a domino effect of errors from which recovery is difficult, and for which manual intervention is inevitable.

Our goal is to remedy this problem, and thus eliminate a major bottleneck in an otherwise automated flow of tomogram generation for cryo-ET. Rather than model the underlying projection geometry for corresponding one marker at a time, we infer 2D geometric similarities between groups of markers in neighboring images. For example, in addition to searching for a single matching marker in an adjacent image, we also search for a matching group of markers that has a similar spatial arrangement in an adjacent image.

While using this method increases the accuracy of resulting correspondence between markers, any system is likely to still make errors. To build an alignment system that is as automatic as possible also requires making the projection model estimation algorithm robust to any remaining correspondence errors. We present a probabilistic framework which combines all of the above goals in order to automatically generate a full-precision aligned tilt series.

The alignment of images using RAPTOR involves the steps presented in Fig. 1. The only user intervention is to select 3 markers to have a template of what a marker looks like in each particular dataset. Using the normalized cross-correlation (NCC) between the template and an image we find possible marker locations in each image of the tilt ser-

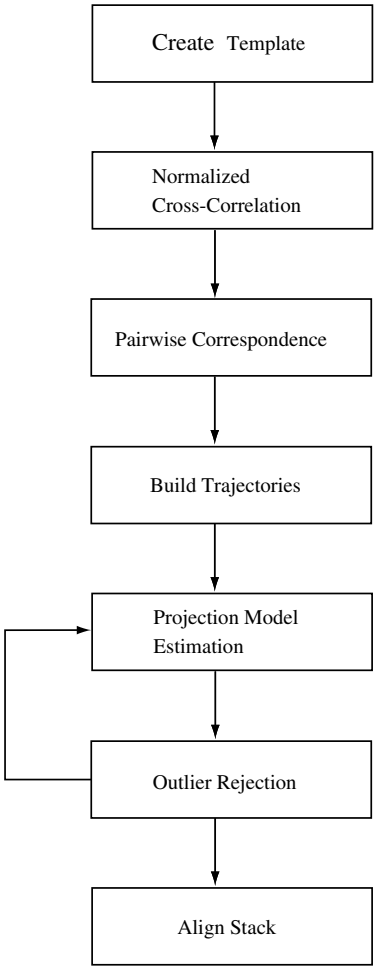

Fig. 1. Block diagram indicating the basic steps to align a tilt series. First, we identify possible markers in the image by creating a template for the markers, and then generating a cross-correlation map of the image and the template. Peaks in that map are possible marker locations. Second, using maximum likelihood estimation, we find correspondence between markers in adjacent and nearby images and use this information to estimate the trajectories of the markers. Finally, we fit a global projective model to align the tilt series. This block diagram is common, in general terms, to all the existing alignment software.

Please cite this article in press as: Amat, F. et al., Markov random field based automatic image alignment ..., J. Struct. Biol. (2007), doi:10.1016/j.jsb.2007.07.007

ies. If the user selects to target M markers per image, between $3M$ and $4M$ peaks are selected in each image. Once we have candidates in each image we find pair-wise correspondence between them. Combining pair-wise correspondences we obtain full trajectories. Once we have trajectories we try to fit a linear projection model using robust optimization to automatically reject outliers and estimate alignment parameters. Even though RAPTOR was initially targeted for low dose cryo-ET images, we show in Section 4 that it also works well in other scenarios such as plastic-embedded sections and X-ray tomography.

Section 3 describes the overall framework, and explains in more detail our methodology for local correspondence, global correspondence, and projection model estimation. Having described our algorithm, Section 4 then presents results using challenging datasets that previously required manual intervention. Section 5 discusses our results in the light of previous work; finally, we summarize the case for this new algorithm in Section 6.

2. Materials and methods

The datasets used to test RAPTOR consisted of: 8 *Caulobacter crescentus* cryo-ET tilt series acquired by hand; 2 *C. crescentus* cryo-ET tilt series acquired automatically using the package UCSF Tomo, provided to us by Grant Jensen from Caltech; 4 *Deinoccocus grandis* cryo-ET tilt series acquired by hand; 2 bacteriophage Phi29 cryo-ET tilt series acquired by hand; one automatically acquired tilt series of a human Eukaryotic cell plastic section sample (Golgi) labeled with quantum dots, provided to us by Mark Ellisman from UCSD; two automatically acquired tilt series of a human Eukaryotic cell plastic section sample with high density of gold markers. Finally, we used 4 X-ray datasets obtained from Carolyn Larabell at the National Center for X-ray Tomography at Lawrence Berkeley National Laboratory (LBNL). These datasets were chosen from a much larger dataset acquired by one of the authors (LRC), according to the specific challenges they posed. Most datasets chosen are of worse quality or posed more difficult challenges than standard tilt series, such as very low SNR, jittery motion, incomplete images, or very thick cells. The most meaningful results are reported in Table 1.

3. Probabilistic image correspondence and projection estimation

We seek the most likely projection model given the observed 2D image data. To find this model involves detecting marker locations in each image, and then recovering the marker trajectories across the tilt series, and finally fitting a projection model to the estimated trajectories. The unknowns of the problem are therefore the 3D marker locations, the 2D marker locations in each image, and the projection model. Let $\mathbf{R}$ be a matrix containing the 3D marker locations, and $\mathbf{S}$ be a matrix containing the true 2D locations of the markers in each image. Finally, let $\mathbf{O}$ to be the noisy measured values of $\mathbf{S}$. The vector of parameters $\mathbf{\Theta}$ is the projection model that relates $\mathbf{R}$ to $\mathbf{S}$. Finding the best projection model is now the same as finding the arguments which maximize the probability of $\mathbf{R}$, $\mathbf{S}$, and $\mathbf{\Theta}$ given the observed data $\mathbf{O}$. This probability can be expressed as the product of two terms:

$$P(\mathbf{R},\mathbf{S},\mathbf{\Theta}|\mathbf{O}) = P(\mathbf{R},\mathbf{\Theta}|\mathbf{S},\mathbf{O})P(\mathbf{S}|\mathbf{O}) \quad (1)$$

The right-most term is basically the solution to the correspondence problem, estimating the probabilities of the ordered marker positions given the image data. The preceding term gives the probability of a projection model, given the correspondence.[2]

We now describe our approach in more detail. Section 3.1 describes the first step, detecting markers. In Section 3.2 we discuss how all markers in two images are used to find correspondence for that pair of images. Section 3.3 describes how we use multiple pair-wise correspondences to estimate $\mathbf{S}$, or the marker trajectories. Finally, Section 3.4 describes how we find the most likely projection model once $\mathbf{S}$ is known.

3.1. Marker detection

The very first step is to detect markers in all the images, and to provide a score for each detected marker. For this step, a template is required. The user picks a small number (we found 3 to be sufficient) of marker centers, and the program then averages the corresponding marker patches. This provides a preliminary template. We then cross-correlate this template with all the images at tilt angles between $-30°$ and $+30°$, do a peak search and select the first five peaks from each of those images. The particles corresponding to these peaks are then averaged to obtain a template that is used for the final cross-correlation. The noise of the new template is much reduced relative to the preliminary template. The selection of the first three markers for the preliminary template is the only manual step required in RAPTOR, and is much faster than a standard initial seeding step used in many packages. Fig. 2(A) and (B) show a comparison between an individual (of three) marker used for the preliminary template and the final template generated by this approach.

In a last pass, we once again cross-correlate the final template with all the images, do a peak search, obtain all the peaks and apply a threshold. The resulting set of peaks in each image is our list of candidates in that image. If the number of target markers is M the number of peaks selected as candidates per image ranges from $3M$ to $4M$.

3.2. Probabilistic correspondence for one pair of images

Corresponding groups of markers in two images amounts to assigning each of the markers in a group in

[2] This model is described in more detail in the Supplementary materials.

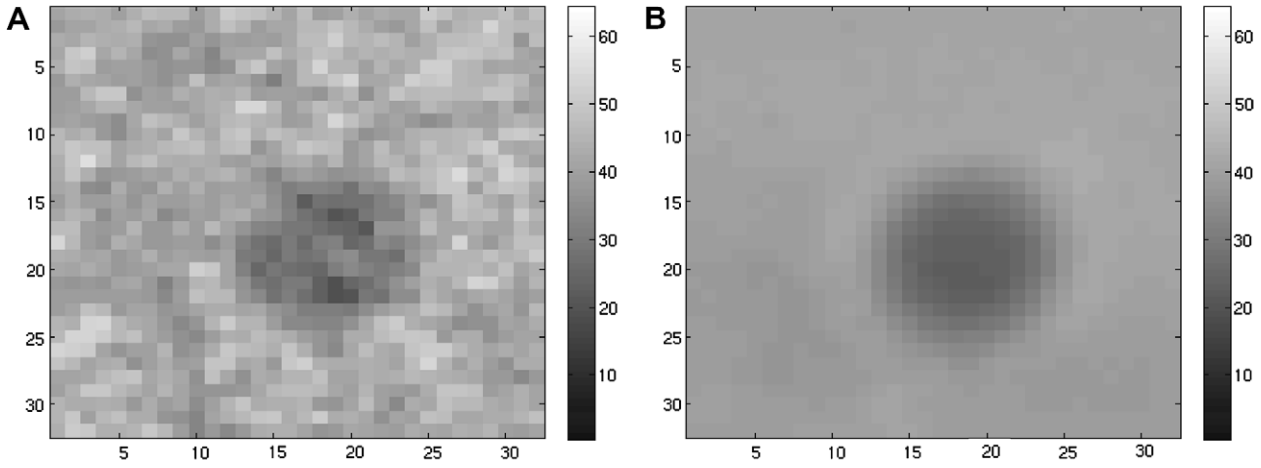

Fig. 2. A gold marker and the final template for one cryo-EM tilt series. Three initial markers as the one shown in (A) are chosen by the user. RAPTOR uses them as references to create a final template, shown in (B). The template is refined automatically selecting the highest peaks of the normalized cross-correlation procedure in each image at low tilt angle.

one image to one (or none) of the markers of the other image. This is analogous to mapping stars in two images of the night sky (Fig. 3). At one extreme, we could attempt to solve the correspondence of all the stars simultaneously. This approach is prohibitively complicated. Alternatively, locating one star in isolation leaves much room for error (Fig. 3(A)). A good compromise solution is to break the sky up into constellations, locate each one for every image, and then to build the global map based on these local mappings (Fig. 3(B)—it is difficult to mistakenly find another "big dipper"). Intuitively, once given a star's constellation, there is little information to be gained about that star from all the other stars in the sky. This intuition is captured mathematically in the notion of conditional independence, which is the underlying mechanism of our method.

To employ this intuition in a probabilistic framework, we must encode our preferences for what constitutes a likely solution into an overall joint distribution for all the possible correspondences between markers. For example, if we believe a group of markers should form the shape of the big dipper, any group of candidate markers that form such a shape would receive a higher weight.

In order to utilize information from more markers in each decision, we define a distribution over the assignments of the markers of one image to another. Specifically, let $\mathbf{A} = \{A_1, \ldots, A_M\}$ be the set of markers in the first image and let $\mathbf{B} = \{B_1, \ldots, B_K\}$ be the set of markers in the second image. Also, let I_1 and I_2 be the two images themselves. We now treat each A_i as a random variable that can take a value in the set $\mathbf{B}$ in addition to an 'unassigned' value. The joint distribution

$$P(A_1, \ldots, A_M | I_1, I_2)$$

assigns a probability value for each assignment of the markers in A. Our goal is to find the assignment that maximizes this distribution,

$$A^* = \operatorname{argmax} P(A_1, \ldots, A_M | I_1, I_2) \quad (2)$$

The complexity of this operation is in general exponential in the number of markers M. However, if the distribution P can be factorized, the complexity can be reduced significantly (to polynomial). Factorized distributions arise when the variables have conditional independence. This is a reasonable assumption in our case, as explained above.

A set of conditionally independent random variables (that is, they depend on each other only through subsets of those variables) is called a Markov Random Field (MRF).[3] In our formulation, we choose proximity between two markers as a criterion for dependence. That is, we define a pair of markers to depend on each directly if they are within some distance of each other. A graphical example of such a group of six markers is shown in Fig. 3(D). An edge between two markers indicates direct dependency. As mentioned previously, the joint distribution of an MRF[4] factors into local terms, each of which encodes our preferences for those localities. To further contain computational complexity, we limit our factors to be over at most two markers.

Even with the simplified distribution, exact inference of marginal and conditional probabilities needed for correspondence is not practical. *Loopy belief propagation* (LBP) is an approximate inference method that has been shown to converge to good results for MRF problems in imaging applications (Murphy and Weiss, 1999). For a detailed description of LBP the reader is referred to (Yedidia et al., 2000, 2005; Kschiang et al., 2001; Bishop, 2006).

[3] The theory of MRF's is very rich and beyond the scope of this paper. For a thorough treatment, the reader is referred to the many references on the subject (e.g. Yedidia et al., 2005; Bishop, 2006). We provide a brief discussion as part of the overall probabilistic framework in Supplementary materials.

[4] In RAPTOR, the threshold is one quarter of the image (e.g. two markers closer than 512 pixels in a $2K \times 2K$ image will be connected in the graph). Therefore, each marker is connected to many other markers. In addition, we impose that each marker has a minimum of 8 neighbors, even if it is necessary to violate the threshold.

Please cite this article in press as: Amat, F. et al., Markov random field based automatic image alignment ..., J. Struct. Biol. (2007), doi:10.1016/j.jsb.2007.07.007

Below we provide a brief example to illustrate its basic idea. In RAPTOR we use a variant of LBP (Elidan et al., 2006) that has fast convergence properties.

LBP attempts to infer the correspondence by combining local information obtained from the factors of the distribution. It does this by propagating each piece of local information through the MRF to achieve a globally consistent assignment between markers and candidates. An artificially constructed example of how this might work is shown in Fig. 3(D–G). A reference image (Fig. 3(D)) has six markers (red, blue, green, purple, brown, and cyan) to be corresponded in a candidate image (Fig. 3(E)). For each reference marker, our "beliefs" on possible candidates are taken from a region around the location of that marker in the candidate image. Note that even for such a small set (6 markers) the number of possible final solutions is quite large (2430). Rather than examining all of these, we infer a selection by considering smaller groups of pairs of candidates at a time. A comparison of each marker and its neighbors in the reference image shows the directions along which corresponding candidates should lie. Applying this constraint in the candidate image narrows down the potential correspondence. In one first iteration, the blue candidates are narrowed down by comparing orientations with respect to the red and green candidates with corresponding orientations of the respective markers. Simultaneously, all other candidates are similarly narrowed down by comparing orientations with the candidates of neighboring markers (Fig. 3(F)). Continuing this procedure, one more iteration completely narrows down the choices to the correct solution (Fig. 3(G)). It should be noted that all comparisons between neighbors happen simultaneously in one iteration by looking at the narrowed-down candidates of the previous iteration. In practical datasets such obvious hints do not typically exist, and more iterations are required for the solution to converge.

In the actual algorithm, we infer the probability from a set of factors (also known as potentials) Φ_{s0} and Φ_{p0}. Φ_{s0} is called the singleton potential, and encodes how likely the candidate marker matches the current target using information only about the target. Φ_{p0}, the pair-wise potential, encodes how likely it is that a pair of close candidate markers match a pair of markers in the original image.

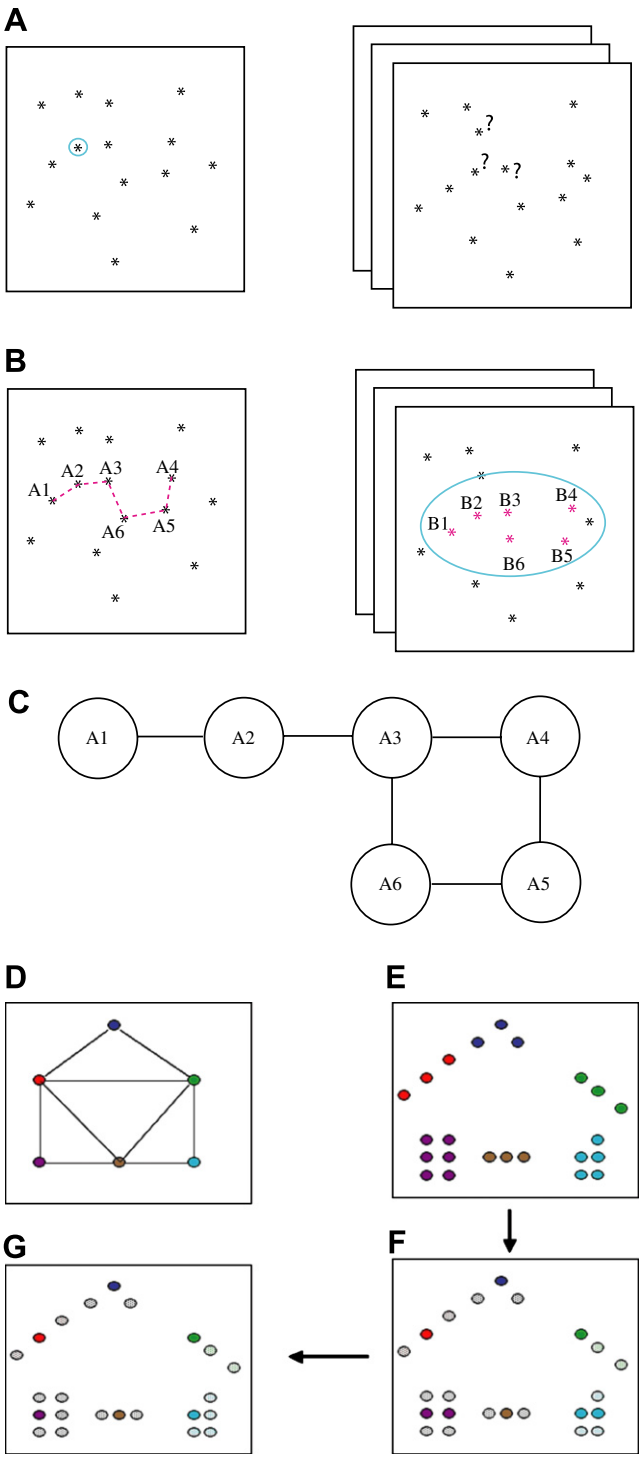

Fig. 3. Local correspondence for one pair of images. (A), Finding the correspondence by looking at a single marker is difficult since it requires identifying a marker in the reference image and choosing one of several (quite similar) candidates in a nearby neighborhood in the candidate image. (B) On the other hand, if a group of markers is considered at a time, the contextual information (e.g. the "Big Dipper" pattern) can be exploited to find a more reliable correspondence. (C) To represent this context when finding correspondence, we represent local relationships as a dependency graph for MRFs. One possible dependency graph is shown here for the MRF formed by the 6 markers in panel (B). Panels (D–G) illustrate qualitatively how candidates for a group of 6 markers are selected in loopy belief propagation (LBP). The 6 markers are shown in panel (D). Markers in two subsequent images are considered neighbors if they are closer than a threshold distance to each other. Each marker has several candidates shown in panel (E). In each iteration of LBP, candidates for each marker compare their orientations relative to candidates for the neighbors of that marker. The highest scoring candidates are the ones that have the most consistent set of such orientations. For example, from panel (E) to panel (F), it is clear that the most likely blue marker candidate is the middle one, since it maintains a consistent orientation with both the red and green marker candidates. A similar set of winners emerges in the purple and cyan group when comparing orientations with the brown group. Repeating this process for the remaining groups of candidates gives us similar winners, which are then compared in the next iteration of LBP. In this carefully constructed case, LBP converged to the right correspondence in two iterations.

Please cite this article in press as: Amat, F. et al., Markov random field based automatic image alignment ..., J. Struct. Biol. (2007), doi:10.1016/j.jsb.2007.07.007

More formally, we assume in each pair of images that there is a reference image (used for conditioning) and a candidate image. There is a list of M corresponded markers A_i ($i = 1 \ldots M$) in the reference image, and another list of K candidate markers B_j ($j = 1 \ldots K$) in the candidate image (A and B are in $\mathscr{R}^2$). Typically, K is taken to be 3–4 times larger than M. The markers A_i, can be either from raw or prealigned images, and can also incorporate any coarse projection model information if that is available.

The singleton potential is based on physical similarity and proximity, and is calculated by cross-correlating a square patch of fixed size centered at a marker extracted from the reference image with the same size patch centered at the candidate marker j. This score is then multiplied by an exponential roll-off factor which is a function of distance from the center of marker A_i to candidate marker B_j. For unaligned images the roll-off is quite mild, with decay constant of about 1/3 the image width.

We define f_k as this exponential fall-off with distance:

$$f_k(\vec{a}, \vec{b}) = \exp\left\{ -\left(\frac{||\vec{a} - \vec{b}||}{k} \right)^2 \right\} \tag{3}$$

This term is a weighting function which decays exponentially with the distance between any pair of vectors (a and b above).

The singleton potential becomes:

$$\psi_{s0}(i_1, j_1) = \mathrm{NCC}(i_1, j_1) \cdot f_{k_1}(\vec{x}_i, \vec{x}_j) \tag{4}$$

where NCC is the normalized cross-correlation of the two markers, and x_i is the position vector of the marker. The pair-wise potential is based on relative orientation and proximity. Let the vector $v_i = A_{i1} - A_{i2}$, and the vector $v_j = B_{j1} - B_{j2}$. The difference vector, $v_i - v_j$ is a measure of the relative motion between marker $i1$ and marker $i2$ between the reference image and the candidate image. The pair-wise potential is the product of three exponentials: an exponential roll-off for distance A_{i1} to B_{j1}, another exponential roll-off for distance A_{i2} to B_{j2}, and finally another exponential roll-off for the L_2 (Euclidean) norm of the difference vector:

$$\psi_{p0}(i_1, i_2, j_1, j_2) = f_{k_2}(\vec{v}_i, \vec{v}_j) \cdot f_{k_1}(\vec{x}_{i_1}, \vec{x}_{j_1}) \cdot f_{k_1}(\vec{x}_{i_2}, \vec{x}_{j_2}) \tag{5}$$

Here the values of k_1 are fairly large, while the value of k_2 is much smaller.

It should be noted that the geometric constraints are applicable to different scenarios, since they make no assumptions about the projection model. They only assume that the there are no big changes in relative marker positions between consecutive projections. This assumption is valid for most projection models used in tomography. At the same time, they include relative position and direction between markers.

3.3. Global correspondence (multi-level approach)

While the pair-wise correspondence now makes use of all the markers in the two images, it still only considers two images at a time. To more robustly generate the marker trajectories across the tilt series, **S**, we compare each image with several of its following images. We use the redundant information both to allow trajectories to skip over a missing marker in some images, and to reduce the probability that a bad correspondence will affect the global trajectories.

The first step of global correspondence is to build trajectories. To do this we use a multi-level approach. The first level contains correspondences between the jth image and the $j + 1$th image. The second level contains the correspondences between the jth and the $j + 2$th image. We use up to 3 levels in images spaced up to 2° apart because after that it is difficult to find correspondences due to the tilt difference between projections. We initialize the first trajectory as the first point in image 1 and level 1. We call this point p_{11} (in general, we define p_{ij} as the point in trajectory i and image j). Local correspondence gives us the location of this trajectory in image 2, i.e. p_{12}. Next, we try to find p_{12} in the local correspondence between image 2 and image 3. If we find it, that returns p_{13} and we repeat the process for this new location to find p_{14}.

We use the extra levels in two ways. First, if a local correspondence for a track is missing, we look in the next level of the global correspondence to complete the track. For example, if we cannot find p_{13} from p_{12} in the first level, we try to find p_{14} from p_{12} in the second level, which contains local correspondence between image 2 and image 4. If we find p_{14}, we continue the process for this new location to try to extend the trajectory. If we do not find it, we try to find p_{15} from p_{12} in the third level, which contains local correspondence between image 2 and image 5. We iterate this process until the first trajectory can not be extended longer, either because we reach the last image (full trajectory) or because one location can not be found in any level (partial contour). To create new trajectories, we use points in each level that have not been used in previous trajectories, since they are potential seeds of new trajectories.

Having multiple match levels allows trajectories to recover from occlusion. Occlusion occurs if a marker present in one image can not be found, even by a human eye, in the next image. A variety of events can create occlusion. For example, a fiducial marker in the edge in one image can move out of the field of view in the next image. Or some other feature in the sample can occlude the marker in some views (two clouds of markers crossing each other, as shown in Fig. 4(A–C)).

The second way of using a multi-level approach, and the second step of global correspondence, is to add confidence in the trajectories obtained in the first step. We do this using the redundant information contained in the local correspondences that have not been used in the first step. For example, suppose we found p_{12}, p_{13}, and p_{14} using the first step. We can check this path against the local correspondence for p_{12} in image 4, which is contained in the second level. We count the number of times such comparisons fail in each trajectory. If this occurs more than one fifth of the

Please cite this article in press as: Amat, F. et al., Markov random field based automatic image alignment ..., J. Struct. Biol. (2007), doi:10.1016/j.jsb.2007.07.007

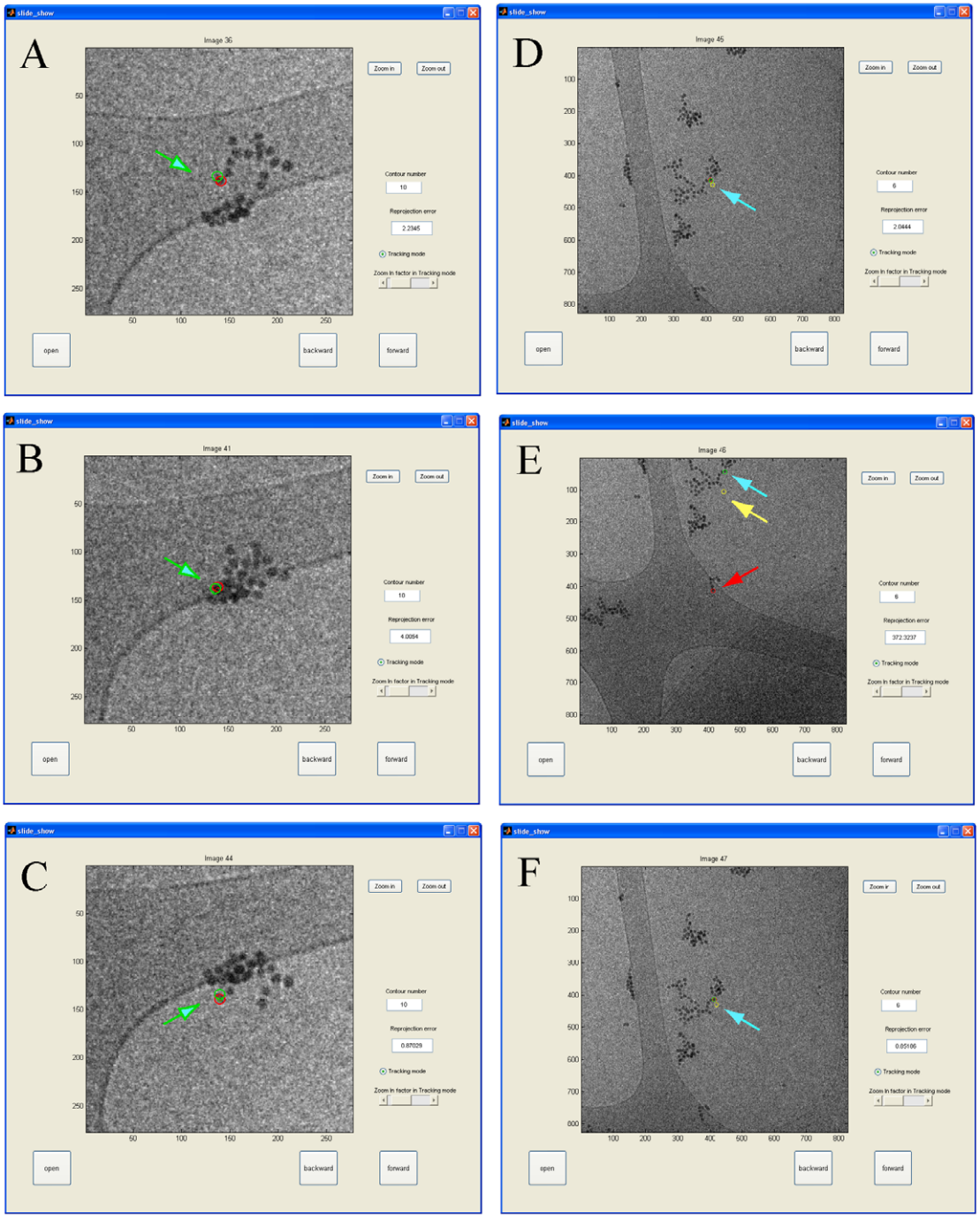

Fig. 4. Robustness of the local and global correspondence. All images are magnified sections from images in the actual tilt series. (A–C) show the tracking of a chosen marker whose trajectory in 2D projections passes through a group of similar markers throughout the tilt series. This situation is particularly difficult for methods which track single markers at a time. The 2D projections (A–C) are spaced by angular increments of more than 6°, for better visualization. In this figure, the red circles indicate the markers chosen by tracking, the green circles are the estimated marker positions using the Huber penalty, and the yellow circles (when shown) are estimated marker positions using a least squares optimization. Panel (D) shows the position of a chosen marker whose trajectory is lost in (E), and subsequently recovered in (F). Panel (E) shows that the least squares estimate is off when a tracking error occurs while the use of the Huber penalty is able to tolerate this error.

length of this particular trajectory, we drop the trajectory. Otherwise, we retain the consistent part of the trajectory.

Finally, only trajectories of certain length are kept to estimate the projection model. The length is selected automatically to assure that we have enough points in each image to estimate the projection model. As explained in the next section, the projection estimation is able to handle partial trajectories of any length across the tilt series. This is extremely useful at high tilt angle, where many markers are present in only a few images before disappearing.

Please cite this article in press as: Amat, F. et al., Markov random field based automatic image alignment ..., J. Struct. Biol. (2007), doi:10.1016/j.jsb.2007.07.007

3.4. Robust estimation

Once we have recovered trajectories across the whole tilt series, a linear projection model can be estimated to represent the transformation between the 3D location of the markers and their 2D projection in each image. The projection model finds a common origin of coordinates for all the images in order to align them. Details about projection models can be found in (Lawrence et al., 2006).

We use convex optimization techniques (Boyd and Vandenberghe, 2004) to estimate a linear projection model. For each jth image in the tilt series and ith marker trajectory we have the following projection model:

$$\begin{bmatrix} x_{ij} \\ y_{ij} \end{bmatrix} = [G_j] \begin{bmatrix} X_i \\ Y_i \\ Z_i \end{bmatrix} - \begin{bmatrix} tx_j \\ ty_j \end{bmatrix} \tag{6}$$

$[x_{ij}\ y_{ij}]^{\mathrm{T}}$ are the marker coordinates found by correspondence; G_j is a 2×3 matrix defining an affine transform which models projections of points from 3D to 2D; $[tx_j\ ty_j]^{\mathrm{T}}$ is the translation in 2D of each image and $[X_i\ Y_i\ Z_i]^{\mathrm{T}}$ are the marker coordinates in 3D. To estimate the parameters from the microscope we solve the following minimization problem:

$$\min\left\{\sum_{j=1}^{T}\left(\sum_{i=1}^{M} w_{ij} * L\left(\begin{bmatrix} x_{ij} \\ y_{ij} \end{bmatrix} - [G_j]\begin{bmatrix} X_i \\ Y_i \\ Z_i \end{bmatrix} - \begin{bmatrix} tx_j \\ ty_j \end{bmatrix}\right)\right) + \lambda||G_j - G_{0_j}||_2^2\right\} \tag{7}$$

The unknowns are the parameters of the microscope (G_j, tx_j, ty_j) and the 3D locations of the markers (X_i, Y_i, Z_i). The equation is clearly nonlinear and nonconvex in those unknowns. However, if we fix G, Eq. (7) is linear and convex in the remaining unknowns, and vice versa. We can thus use an iterative scheme to solve only linear equations. First we solve for G, and then we solve for X, Y, Z, and tx, ty. We iterate this process until we converge to a solution. This method has been known for a long time in the optimization literature as alternate convex programming (Ke and Kanade, 2005) and it gives similar performance as nonlinear optimization methods with the advantage that this approach is computationally simpler to solve.

The term w_{ij} in (7) contains the weights for each residual of the reprojection model. This concept was previously used in the work by Brandt and co-workers (2001) to allow the incorporation of partial trajectories in the reprojection model estimation. If a point is missing in the ith trajectory for the jth image, it is assigned a weight $w_{ij} = 0$. Otherwise, we set $w_{ij} = 1$.

The term $\lambda||G_j - G_{0_j}||_2^2$ in (7) is known as a regularization term. It indicates that the matrix G should not be very different from a given matrix G_{0j} which is set by the geometry of the microscope. The parameter λ indicates how much we penalize the fact that G_j deviates from G_{0j}. λ has been tuned using cross-validation with different datasets and it is a fixed parameter.

The matrix G_{0j} that is used as a reference is:

$$G_{0j} = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta_j) & \sin(\theta_j) \end{bmatrix} \tag{8}$$

where α represents the 2D rotation to establish the y-axis as the tilt axis and θ_j is the measured tilt angle of the goniometer. This is a good first order approximation of the microscope's projection model.

Regularization allows the alignment of images even in tilt series where very few markers are present, since it imposes constraints on the shape of G. These constraints indicate that G should be close to a composition of rotations (at least for an ideal microscope). However, it does allow some flexibility to incorporate distortions or changes of magnification between different projections.

Finally, a crucial part of Eq. (7) is the function $L(*)$. This function is known as a cost and/or penalty function. It is a measure of the error between the trajectory given by global correspondence and our reprojection model. The definition of $L(*)$ has a great effect in the estimation of the reprojection model. To the best of our knowledge, all the existing alignment methods use $L(r) = r^2$ (Lawrence et al., 2006; Brandt et al., 2001; Kremer et al., 1996) except for (Brandt and Ziese, 2006). This quadratic error measure establishes that the solution to (7) is the well-known least-squares problem. However, it is also well-known in the optimization literature that least-squares methods are not robust to outliers. Even a single error in a trajectory can bias the entire reprojection model. Thus, a more robust penalty function is used in this work. We use what is known as Huber penalty function (Huber, 1981) in robust statistics. This penalty function uses a combination of linear and quadratic penalties. For small residuals the penalty is quadratic, and small residuals are thus fitted as least squares. Big residuals, however, are penalized using a linear function. This results in large deviations (outliers) not affecting the reprojection estimation model. The fact that large deviations are tolerated and isolated in the linear region makes it easier to detect and correct them automatically without manual intervention. Notice that Huber penalty is still convex, so we avoid the problem of finding a local minimum as a solution to our problem. More details and references about the reprojection model estimation are given in Supplementary information.[5]

4. Results

The algorithm presented in this work was tested using cryo-ET datasets of the bacteria *D. grandis* and *C. crescen-*

[5] The Supplementary materials describe a new technique that combines image segmentation with estimation of the number of segments (or clusters) to refine the centering of the gold beads. This technique is used once we have a good estimation of the marker location. It addresses the problem of finding markers centers when two or more of them collide. Current software identifies the middle point between two markers as the center of each particle, which is clearly wrong.

tus. Each of these datasets poses unique and different challenges. Most of the datasets were acquired by hand, which results in tilt series with more discontinuous trajectories, or "jerky" movements across the angular range than in the case of automated acquisition.

Deinoccocus grandis cryo grids are considerably thick, with the large bacteria forming conspicuous meniscus in the amorphous ice. As a result, projections from high angles usually occlude many of the markers clearly visible at the lower angles. The cryo-grids used with these bacteria are also more irregular, uneven in shapes, contrast, and transparency than for other bacteria more commonly studied by cryo-ET. For this reason, the usual marker selection and tracking applications included in other programs tend to fail. Most of *C. crescentus* and *D. grandis* datasets were acquired with a total dose of at most 100 $e^-/Å^2$, resulting in very noisy individual images. *Caulobacter* cryo-grids are thick and form a meniscus although they are still transparent to the beam. The surface of the grids is not as irregular as with *Deinoccocus* and the high angle occlusion problem is thus less severe. However, in all cases there are, in high tilt angle views, sets of very useful markers towards the image boundaries which disappear at lower angles.

Fig. 2(A) shows one of three markers selected by the user to create a preliminary template. RAPTOR refines the preliminary template automatically selecting the highest peaks of the normalized cross-correlation procedure in each image at low tilt angle.[6] Fig. 2(B) shows the result of the automatic template refinement.

The robustness of RAPTOR predicting marker trajectories is illustrated in Fig. 4. Panel (A) shows the position of a marker that will intersect a whole group as the tilt series proceeds. In panel (B) the projected positions of all markers are tightly clustered. At higher angles, the two groups are again resolved, as shown in panel (C). The algorithm successfully tracks the position of markers in difficult situations such as this one, which can occur quite frequently. The robustness in these situations is a consequence of our incorporation of contextual information in the correspondence. As already pointed out, the number of markers in each neighborhood, which markers and which neighborhoods, do not require user input beyond the initial choice of targeted number of markers per image. RAPTOR is also tolerant to any remaining correspondence errors, and thus successfully avoids propagating errors. Panels (D–F) illustrate the robustness built into RAPTOR using robust convex optimization techniques. Panel (D) shows the position of a chosen marker which is located to the wrong position by global correspondence in a subsequent image, panel (E). At a subsequent step, however, the trajectory is recovered, as shown in panel (F). This example shows that the projection model is not biased by the presence of outliers. This feature allows the automatic removal of outliers and their re-assignment of their correct location. Methods based on least-squares fail in such cases.

We illustrate in Fig. 5 the overall performance of the algorithm in the alignment of a difficult dataset of *C. crescentus*. This dataset was acquired with a helium-cooled stage, and as a consequence the relative position of the markers was significantly more variable than is normally the case with liquid nitrogen-cooled stages (Comolli and Downing, 2005). Panel (A) of Fig. 5 shows the projection acquired at −58°. Panel (B) shows the set of trajectories recovered by RAPTOR. There are complete trajectories across the whole tilt series, there are trajectories that cover the whole angular range of the tilt series but are missing the marker in some images, and there are partial trajectories starting from both ends of the tilt series, disappearing at lower angles. The meniscus formed by the cell and bars of the cry-grid may occlude, at high tilt angles, markers chosen at low angles. At specific angles throughout the tilt series, the overlapped projection of groups of markers which belong to different planes within the cryo-grid may cause gaps in their tracked trajectories. Many gaps also represent locations where the probability of making a correct marker assignment is not sufficiently high.[7] Finally, there are in the high angle projections, many markers with ideal SNR which are left out of the view in the projections at lower angles. All these events are represented in the diagram plotted in Fig. 5(B).

The physical trajectories of all chosen markers in image space (x–y coordinates in pixel number), in the raw data and the aligned data, are shown in Fig. 5(C) and (D). The inset in each panel shows one single trajectory. No pre-alignment of the raw dataset was performed before obtaining the automatically aligned one, as is readily obvious from the cloud shape of marker trajectories in the raw data (Fig. 5(C)). A comparison with the results of manual marker selection and tracking shows indistinguishable final tomographic reconstructions, which is the ultimate performance test.

Table 1 quantitatively compares RAPTOR's performance versus manual alignment by an expert IMOD user in a great variety of datasets. These include cryo-EM datasets as well as plastic-embedded sections and X-ray tomography. The datasets also use different angular steps between projections, as well as different marker diameters and number of markers present in the image. More detail about those datasets is given in Section 2. Quantitative comparison is performed using the noise-compensated leave-one-out (NLOO) method of (Cardone et al., 2005) for both for manual and the automatic reconstruction. RAPTOR resolution is always comparable, or even better in some

[6] Interestingly, RAPTOR performance does not change at all if we use one or the other template because the normalized cross-correlation scores obtained by one or the other are very similar.

[7] We use a high threshold in order to avoid false matches. The ratio between the probabilities of the two most likely assignments must be greater than 2.7.

Please cite this article in press as: Amat, F. et al., Markov random field based automatic image alignment ..., J. Struct. Biol. (2007), doi:10.1016/j.jsb.2007.07.007

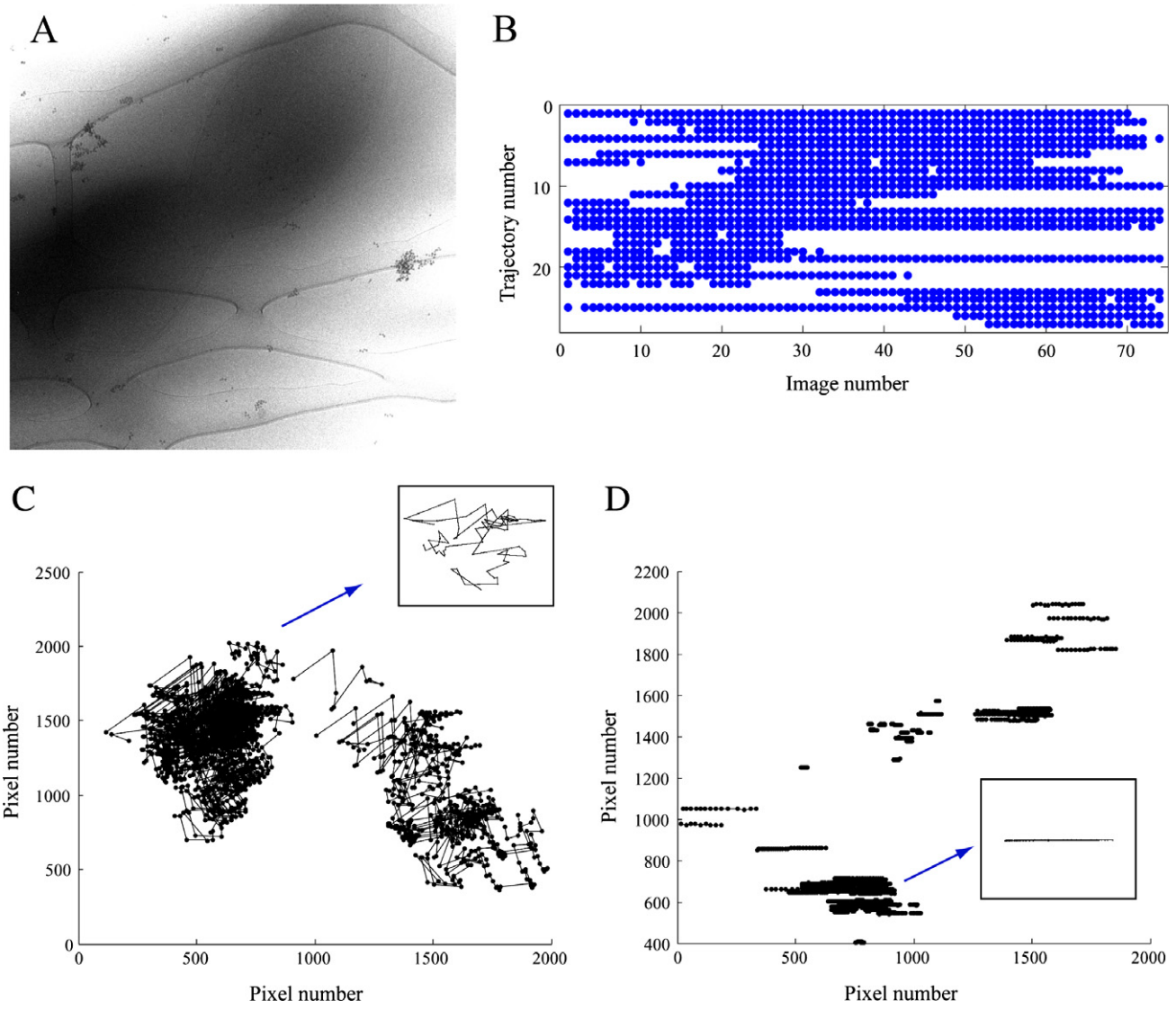

Fig. 5. RAPTOR performance tracking markers in a *Caulobacter crescentus* dataset, and aligning the set. (A) The −58° 2D original projection from the tilt series. The size of the image is 2048 pixels on edge. The cell forms a meniscus of considerable thickness. Both the cell and the grid bars may occlude markers at high tilts. The wider area projected into the images at these angles contains many markers with appropriate SNR, far from the cell, which disappear at lower angles. (B) Trajectories recovered from the tilt series. (C) Overlay of several marker locations in image space (x–y coordinates in pixel number) throughout the raw datasets illustrates the trajectories in the raw dataset, (D) trajectories of same markers after RAPTOR automatic alignment. The insets in (C and D) show the trajectory of a single marker.

datasets, than that obtained by an expert user with extensive manual intervention, demonstrating that RAPTOR obtains full-precision alignment in a variety of difficult scenarios.

The SNR column of Table 1 provides an estimate of the feature visibility in the raw datasets. We chose a signal to noise measure that reflects the visibility of the gold particles in the images as defined in (Frank, 1996):

$$\mathrm{SNR}(\mathrm{gold}_{\mathrm{bead}}) = \frac{\sum_{i,j}|\overline{m}_{ij}|^2}{\sum_{i,j}\sum_{k}|m_{kij} - \overline{m}_{ij}|^2} \tag{9}$$

The measure of the SNR is based on the average of all the gold images. The signal power is the variance of the average gold image, and the noise power is the variance of all the gold images with respect to this average image. m_{kij} is the value of the i,jth pixel in the kth image, and $\langle m_{ij}\rangle$ is the value of the i,jth pixel in the zero mean average image.

For all datasets, we also compared visually the quality of the tomographic reconstructions obtained after RAPTOR alignment with those obtained after alignment by an independent standard method. Two such comparisons are shown in Fig. 6. Panels (A) and (B) show, side by side, a one-pixel-thick slice through a tomographic reconstruction of a dividing *C. crescentus* cell obtained after manual alignment and after RAPTOR automatic alignment, respectively. The quality of both is equivalent. Another

Please cite this article in press as: Amat, F. et al., Markov random field based automatic image alignment ..., J. Struct. Biol. (2007), doi:10.1016/j.jsb.2007.07.007

Table 1
Summary of the datasets tested to asses RAPTOR's performance (Resolution)

Dataset	Tilt range (deg)	Approx. number of markers	Bead diam. (pixels)	SNR (hi tilt)	Number of patches	NLOO3D(0.3) manual (nm)	NLOO3D(0.3) RAPTOR (nm)	NLOO3D(0.5) manual (nm)	NLOO3D(0.5) RAPTOR (nm)
Cryo-Caulo1	−58:2:54	15	10	0.038	200	14.26	13.81	30.16	24.48
Cryo-Caulo2	−34:1:36	10	10	0.026	196	11.02	8.63	15.11	14.25
Cryo-Caulo3	−62:2:58	15	10	0.007	199	18.01	16.35	24.82	26.78
Cryo-Caulo4	−61:2:54	10	10	0.009	184	25.21	32.54	36.59	41.14
Cryo-Caulo5	−60:1:60	70	10	0.294	200	10.35	10.34	13.32	13.18
Cryo-Caulo6	−68:2:68	20	10	0.005	190	17.27	18.68	24.09	27.96
Cryo-Caulo7-He	−54:1.5:51	60	10	0.024	195	15.44	18.34	27.01	27.89
Cryo-Caulo8	−45:1:45	20	10	0.066	195	13.25	13.63	20.01	20.49
Cryo-Caulo9	−60:1:62	10	8	0.032	199	7.51	7.66	14.32	16.95
Cryo-Deino1	−62:2:57	30	10	0.044	200	19.28	19.64	29.13	30.15
Cryo-Deino2	−60:1:57	20	7	0.008	190	14.75	10.56	20.55	15.18
Cryo-Deino3	−25:1:26	40	10	0.044	199	21.04	19.38	30.93	27.65
Cryo-Deino4	−62:2:62	15	10	0.001	196	−1	−1	−1	−1
Plastic1	−58:2:60	>200	4,8	0.471	93	23.924	26.69	42.147	45.42
Plastic2-qdot	−60:2:60	>200	8	1.653	36	19.55	25.45	28.07	32.51
X-ray-EColi1	−90:2:90	30	10	0.095	200	*	*	53.21	52.35
X-ray-EColi2	−90:2:90	25	10	0.050	200	57.14	56.818	71.428	73.652

Columns two to six contain information to present the variety of scenarios where RAPTOR was tested: spacing between projections, approximate number of gold beads to track in the dataset, bead diameter in pixels, SNR of the gold beads in high tilt angles, and number of patches obtained to calculate the SNR. The last four columns show a quantitative comparison of the resolution obtained using human-directed alignment and the automatic full-precision alignment presented in this paper. The method used for resolution assessment is noise-compensated leave-one-out (NLOO) by Cardone et al., using threshold values 0.3 and 0.5. RAPTOR achieves better or equal resolution (lower numbers in the NLOO3D columns) in many of the datasets. The dataset "Cryo-Caulo5" is courtesy of the Grant Jensen laboratory, Caltech. The plastic data sets "Plastic1" and "Plastic2-qdot" were provided by Mark Ellisman's group at UCSD. The group of Carolyn Larabell, LBNL, generously provided us with the X-ray tomographic tilt series listed as "X-ray-EColi1,2".
(*) Curve did not reach that point; (−1) method fails to align.

example is shown in Fig. 6, panels (D) and (E). The features observed in a one pixel-thick slice through a tomographic reconstruction are comparable in definition and resolution, whether the dataset was aligned with human intervention or automatically by RAPTOR. A better comparison is established by careful inspection of regular features, as shown in the insets. These reconstructions were binned by two, from an image size of 2048 pixels on edge to 1024, and each box measures 84 pixels on edge.[8] Fig. 6, panels (C) and (F) show the resolution plots for each final reconstruction obtained using cross-validation available in software Electra (Cardone et al., 2005). The plots demonstrate that RAPTOR obtains the same resolution as the manual reconstruction.[9]

Fig. 7 shows similar comparisons of reconstruction quality with X-ray images of a frozen-hydrated sample of *Escherichia coli* cells and a tomogram of a conventional stained plastic section. Several of the gold fiducials (70 nm diam.) in the X-ray sample are visible in this cross-section of the reconstruction. The high contrast from the capillary initially caused a problem with the marker identification, but this problem was easily solved by applying edge detection techniques and a filter to reduce the sharpness of the capillary edge in the individual images. We have found that RAPTOR has no difficulties with conventional plastic section datasets. However, the total number of markers needs to be kept below an excessive number (fewer than ∼200) in order to make RAPTOR run efficiently, as discussed below.

The next section analyzes RAPTOR's performance in more detail.

5. Discussion

We have shown our application's ability to recover partial marker trajectories from tilt series acquired with jittery motions, low exposures and low SNR. In Fig. 5 we compare the trajectories before and after alignment to show that they form straight lines as if the sample was rotating along the y-axis. Our algorithm, however, does not require pre-alignment of the tilt series as a first step. In all examples shown here we used the raw data as direct input. The trajectories usually form a cloud, representing the jittery motion usually found in a raw tilt series. Our algorithm is robust enough to tolerate jumps between projections of 350 pixels in a $2K \times 2K$ image, and also tolerates a relatively high level of distortion. Nevertheless, if the images are pre-aligned after acquisition, or acquired by efficient automatic acquisition software such as shown in (Zheng et al., 2007), our method runs faster.

[8] Notice that it is impossible to obtain exactly equivalent slices from these reconstructions because the orientation of the reconstructed volume is not identical in data sets aligned by hand or automatically. The quality comparison of the reconstructions seems clear and straightforward nonetheless.

[9] Similar graphs were obtained for each of the datasets in Table 1, only two datasets are shown due to space limitations.

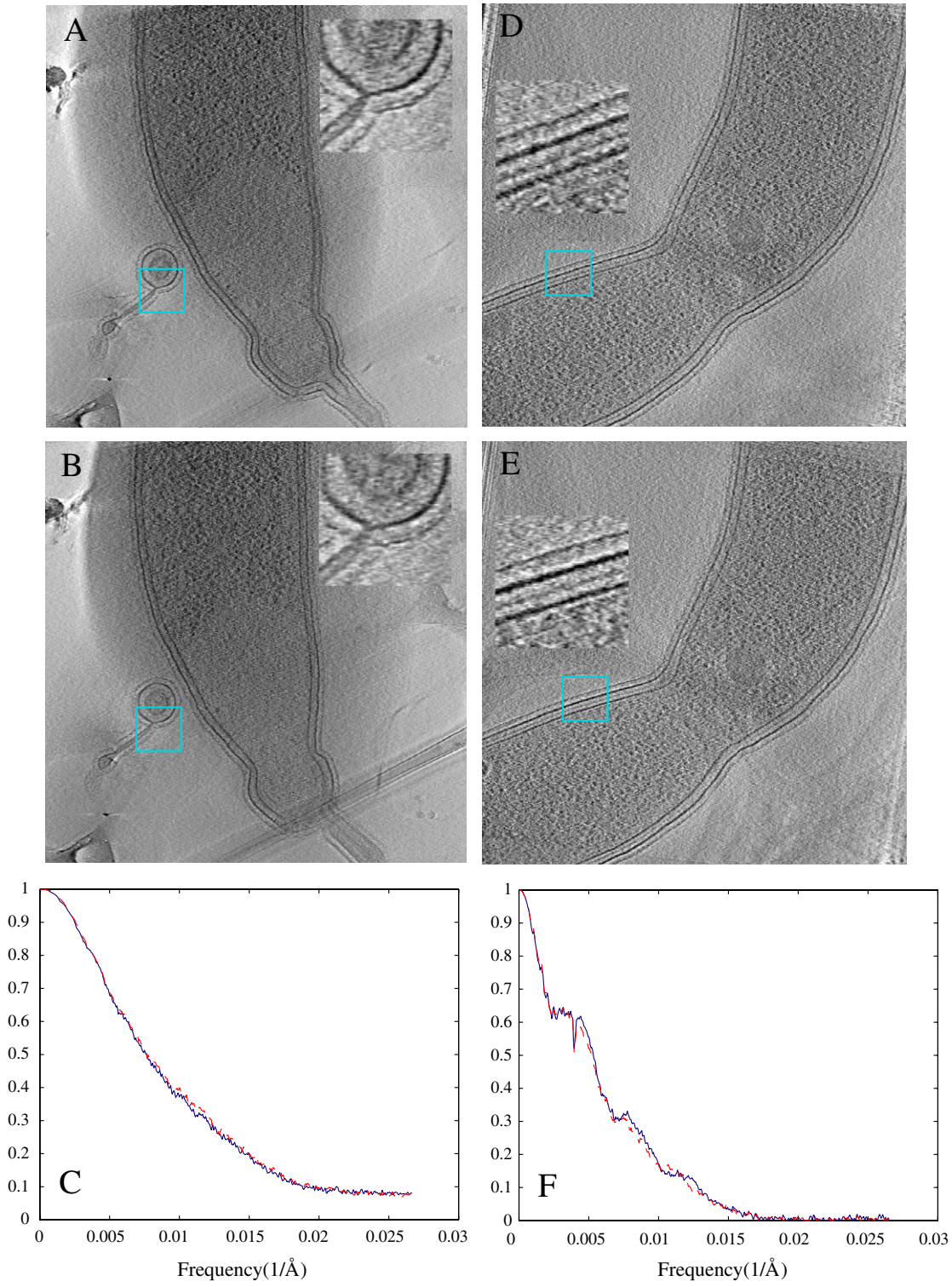

Fig. 6. Tomographic reconstructions based on RAPTOR and manually aligned datasets. (A and B) One pixel-thick slices from tomographic reconstructions of a *Caulobacter crescentus* polar mutant, obtained from a manually aligned dataset and from an automatically aligned dataset, respectively. (D) One pixel-thick slice from a tomographic reconstruction of a dividing *Caulobacter crescentus* cell after manually aligning the tilt series and (E) After RAPTOR automatic alignment. Reconstruction after marker tracking by hand and human-directed alignment (A and B), and after automatic alignment (D and E), results in equivalent final quality. The images are sections, 900 pixels on edge, from tomographic reconstructions binned from 2048 to 1024 pixels on edge. The boxes shown in the insets measure 84 pixels on edge in the binned images. Panels (C and F) show quantitative resolution comparisons for tomographic reconstructions of the two datasets above using noise-compensated leave-one-out method of Cardone et al. The resolution curves obtained with RAPTOR, in red, and with human-directed alignment, in blue, are equivalent. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

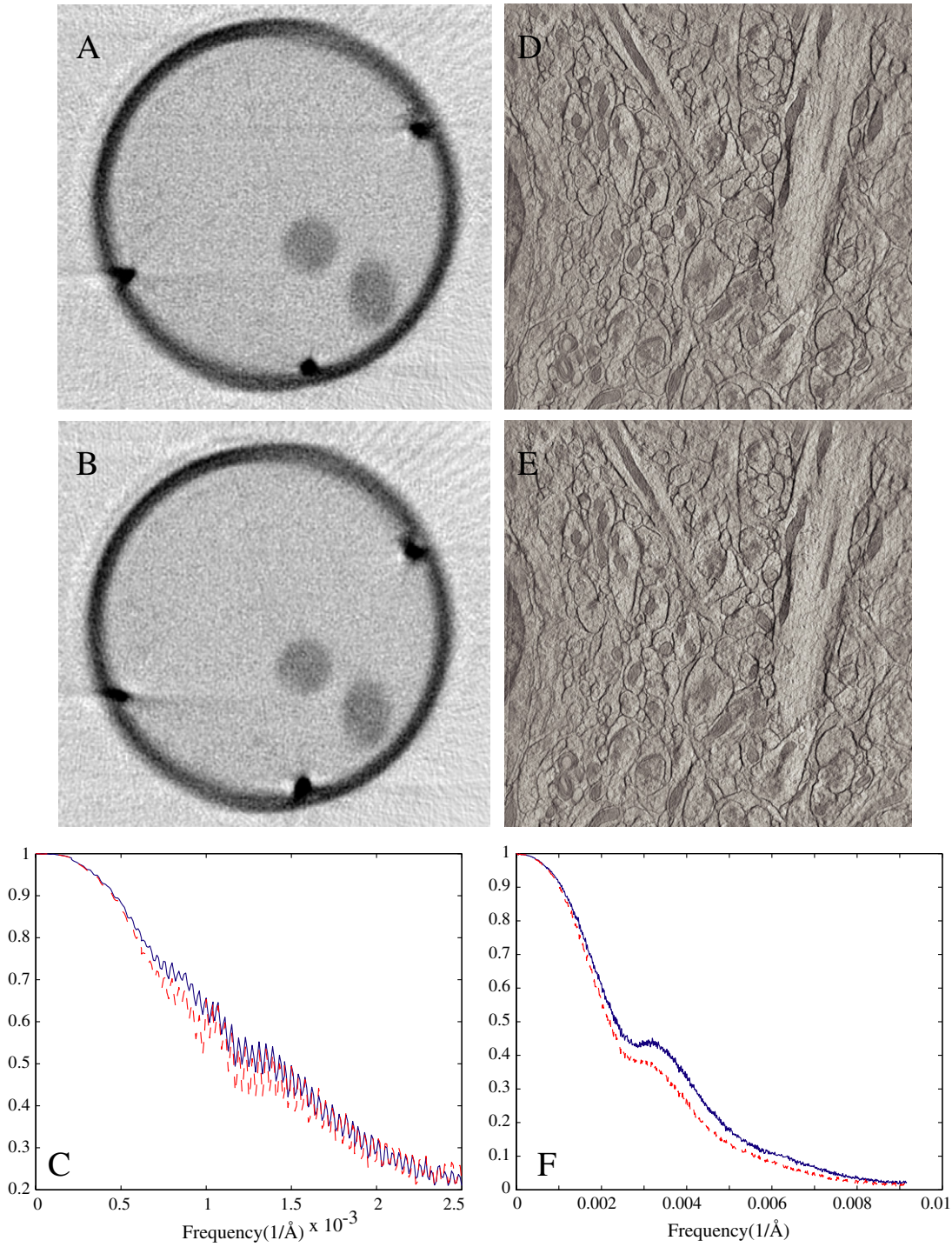

Fig. 7. Tomographic reconstructions from a X-ray tomographic tilt series of intact, frozen-hydrated *E. coli* cells in a glass capillary (A–C), and from a stained plastic section of a eukaryotic cell (D–F). Comparison of results obtained with RAPTOR alignment (A and D), and manual alignment (B and E) show no detectable differences. (A and B) One pixel-thick slices from the X-ray tomographic reconstruction viewed along the tilt axis, showing strong contrast for the capillary with several cells. (D and E) Normal slices through the cellular tomogram. Panels (C and F) show quantitative resolution comparisons for tomographic reconstructions of the two datasets above using noise-compensated leave-one-out method of Cardone et al. The resolution curves obtained with RAPTOR, red, and with human-directed alignment, blue, are equivalent. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

The use of contextual information and simple geometric constraints to establish reliable correspondence has been shown here to be a powerful tool in achieving fully automatic alignment. This was possible due to the random placement of the gold beads, which form unique spatial arrangements between markers. In this way it was possible to build a robust global correspondence that was consistent from image to image even in very noisy datasets. Intuitively

Please cite this article in press as: Amat, F. et al., Markov random field based automatic image alignment ..., J. Struct. Biol. (2007), doi:10.1016/j.jsb.2007.07.007

pair-wise potentials should be adequate because constraining the relative position between pairs of features also constrains the neighborhood as a whole. This allows alignment in various geometry settings as long as the difference between adjacent projections preserves relative distances between markers.

The results summarized in Table 1 show quantitatively how the achieved automatic alignment is comparable to the alignment done using IMOD and manual intervention by an expert user in a very diverse group of datasets. For most datasets the difference in resolution obtained with RAPTOR or manual intervention is within a 3 nm interval. In many cases, RAPTOR obtains a better resolution score. Gold beads of diameter as small as 4 pixels were successfully tracked. The algorithm also adapts automatically to a wide range of marker density in the image.

However, there are three datasets in the table that present some of RAPTOR's limitations. The first one is Cryo-Caulo4. The SNR for the gold beads in the range from $-61°$ to $-55°$ was 0.004 and the NCC was not able to detect the markers. As a consequence a few images were left without being aligned and degraded the resolution of the tomograms as reflected in Table 1. Therefore, our SNR measure for markers tries to show a lower limit for the NCC to find markers. Without identification of possible markers from NCC the rest of the system can not proceed and the alignment fails. This limit seems to be around 0.004 although is not a precise threshold.

In order to further characterize the limits of RAPTOR, we chose from plastic sections datasets with a very high density of gold markers (more than 200 per image). One might think that the high contrast of stained material in those datasets might cause more false matches and break down the pair-wise correspondence step. However, that is not the case. The trajectories obtained with RAPTOR correctly follow the beads as in cryo-EM datasets. The problem in the plastic sections datasets in Table 1 is that the total number of markers per image is very high. It takes too long for RAPTOR to track over 200 markers, so we target 120 markers. With so many markers it is difficult to always have the same markers from image to image in the list of candidates selected using NCC. Therefore, most of the recovered trajectories are far from full trajectories. Partial trajectories make it more difficult to create a precise global alignment which is reflected in a lower resolution score. In plastic, local distortions play an important role and full trajectories are crucial to correct them in the aligned stack. RAPTOR will have an optimal performance with 120 or fewer markers per image.

We have only found one specific case where false matches caused by high contrast areas that are not markers can limit RAPTOR's performance. NCC assigned relatively high score to the boundaries of the cylindrical holder in the X-ray datasets (Fig. 7), which are imaged as straight lines. They create a regular pattern in the peaks across the whole image that affects the inference in the MRF. This problem is easily overcome using standard edge detection techniques followed by a Hough transform to find straight lines. Any peak in that line is erased. RAPTOR can handle a high percentage of false matches in marker detection as long as those false matches have random location in several consecutive images or the structure creating false matches is small enough to not confuse the inference. This is true in most datasets.

One assumption in RAPTOR's local correspondence is that there is a small incremental angle between images in the tilt series. We have successfully tried datasets with incremental angles ranging from 0.5° to 4°. This accommodates most common tomographic tilt series. However, for bigger angular increments, two problems will arise. First, adjacent images will be less similar. This might cause a decrease in the number of correspondences found between images. Second, it becomes more difficult to have many levels in the multi-level scheme, which might affect the robustness of the method. Although we can address this problem by modifying the global correspondence scheme it has not been necessary for any datasets that were available to us.

While an expert user needed up to a full a day to completely align some of our datasets to sub-pixel accuracy, our method is able to align the images automatically. RAPTOR takes 45 min to align a set of 65 $2K \times 2K$ images with 20 markers, and roughly 4 h to align a set of 80 images, ($2K \times 2K$ in size), targeting 70 markers per image in a 2.4 GHz Pentium 4 PC with 2 GB of RAM. It takes roughly 7 h for 240 images ($2K \times 2K$ in size) targeting 30 markers per image. The algorithm is highly parallelizable, so the time can be greatly reduced even using a small cluster available in most cryo-EM labs. It is worth mentioning that the execution time strongly depends on the number of markers present in each image.

The goal of this initial version of the program was to keep it as simple and automated as possible, and make its data files compatible with IMOD. In this way the user can use IMOD to manipulate the data before or after the alignment. While having a turn-key solution to image alignment is useful, it should be possible to extend this tool to make it more flexible, and provide a little more user control. For example, RAPTOR already allows the user to seed specific markers, or to provide multiple templates to track different kinds of features in each dataset. This is very helpful in plastic sections where two sizes of markers are used to distinguished top from bottom. These enable its use on a wider class of images. However, more extensions (e.g. dual axis support) can be easily added in the future.

6. Conclusions

A new probabilistic framework for automated full-precision alignment of EM tilt series images has been presented. We tested our algorithm on datasets where we had manual

Please cite this article in press as: Amat, F. et al., Markov random field based automatic image alignment ..., J. Struct. Biol. (2007), doi:10.1016/j.jsb.2007.07.007

full-precision alignment and reconstructions available, and the results had the same resolution, using visual and quantitative metrics (Cardone et al., 2005). The underlying probabilistic framework allows RAPTOR to align these difficult datasets.

RAPTOR has two main advantages. First, the framework of Markov Random Fields enables a more reliable correspondence through the use of contextual information. The correspondence can be established for any geometry as long as the tilt difference between projections is small enough to preserve relative distances between markers. Second, the use of robust optimization and redundancy in the global correspondence enable automatic recovery from errors, which are difficult to avoid in such noisy environments. By making use of these powerful, yet efficient techniques, RAPTOR is able to successfully complete the task on a relatively common desktop machine in a few hours, enabling high throughput generation of full-precision tomograms.

RAPTOR is freely available for academic use at http://www-vlsi.stanford.edu/TEM/index.htm. Interested users can contact Farshid Moussavi at farshid1@stanford.edu and Fernando Amat at famat@stanford.edu for more information.

Acknowledgments

We thank Daphne Koller of the Stanford University Computer Science Department for her tremendous guidance. We also thank Geremy Heitz from the Stanford Computer Science Department for the many fruitful discussions and important support for the implementation of our correspondence methods. F.M., F.A., and M.H. give special thanks to Harley McAdams, Department of Developmental Biology, Stanford University School of Medicine, for his moral and financial support, without which this work would not have been possible.

We also thank the National Center for Microscopy and Imaging Research (NCMIR) at UCSD for all the beta testing on RAPTOR and the plastic datasets. Also many thanks to Carolyn Larabell and her team at the National Center for X-ray Tomography (NCXT) for the X-ray datasets. They are funded by the National Center for Research Resources of the National Institutes of Health (P41 RR019664-02). We thank Grant Jensen's lab in Caltech for cryo-EM datasets using automatic acquisition software , and Grant Bowman and Lucy Shapiro for the use of their unpublished Caulobacter image in Fig. 6. Finally Fernando Amat thanks Argyris Zymnis of the Stanford University Electrical Engineering Department for his help in coding the optimization part in RAPTOR.

This work was supported by the Director, Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

Appendix A. Supplementary data

Supplementary data associated with this article can be found in the online version, at doi:10.1016/j.jsb.2007.07.007.

References

Baumeister, W., 2002. Electron tomography: towards visualizing the molecular organization of the cytoplasm. Curr. Opin. Struct. Biol. 12, 679–684.

Bishop, C.M., 2006. Pattern Recognition and Machine Learning. Springer, Heidelberg, Germany.

Boyd, S., Vandenberghe, L., 2004. Convex Optimization. Cambridge University Press, Cambridge, UK.

Brandt, S., Heikkonen, J., Engelhardt, P., 2001. Multiphase method for automatic alignment of transmission electron microscope images using markers. J. Struct. Biol. 133, 10–22.

Brandt, S., Ziese, U., 2006. Automatic TEM image alignment by trifocal geometry. J. Microscopy 222, 1–14.

Cardone, G., Grünewald, K., Steven, A.C., 2005. A resolution criterion for electron tomography based on cross-validation. J. Struct. Biol. 151, 117–129.

Comolli, L.R., Downing, K.H., 2005. Dose tolerance at helium and nitrogen temperatures for whole cell electron tomography. J. Struct. Biol. 152, 149–156.

Elidan, G., Mcgraw, I., Koller, D., 2006. Residual belief propagation: informed scheduling for asynchronous message passing. Proceedings of the 22nd Conference on Uncertainty in Artificial Intelligence.

Frank, J., 2006. Electron Tomography, second ed. Plenum Publishing Corporation, New York.

Frank, J., 1996. Three Dimensional Electron Microscopy of Macromolecular Assemblies. Academic Press, Inc., San Diego.

Heymann, J.B., 2001. Bsoft: image and molecular processing in electron microscopy. J. Struct. Biol. 133, 156–169.

Huber, P.J., 1981. Robust Statistics. John Wiley & Sons, New York, NY, USA.

Jensen, G.J., Briegel, A., 2007. How electron cryotomography is opening a new window onto prokaryotic ultrastructure. Curr. Opin. Struct. Biol. 2, 260–267.

Ke, Q., Kanade, T., 2005. Robust L1 norm factorization in the presence of outliers and missing data by alternative convex programming. IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2005), San Diego, CA.

Kschiang, F.R., Frey, B.J., Loeliger, H., 2001. Factor graphs and the sum-product algorithm. IEEE Trans. Inf. Theory 47, 498–519.

Koster, A.J., Grimm, R., Typke, D., Hegerl, R., Stoschek, A., Walz, J., Baumeister, W., 1997. Perspectives of molecular and cellular electron tomography. J. Struct. Biol. 120, 276–308.

Koster, A.J., Klumperman, J., 2003. Electron microscopy in cell biology: integrating structure and function. Nat. Rev. Mol. Cell Biol. 4, SS6–SS10.

Kremer, J.R., Mastronarde, D.N., McIntosh, J.R., 1996. Computer visualization of three-dimensional image data using IMOD. J. Struct. Biol. 116, 71–76.

Lawrence, A., Bouwer, J.C., Perkins, G., Ellisman, M.H., 2006. Transform-based backprojection for volume reconstruction of large format electron microscope tilt series. J. Struct. Biol. 154, 144–167.

Mastronarde, D.N., 1997. Dual-axis tomography: an approach with alignment methods that preserve resolution. J. Struct. Biol. 120, 343–352.

Medalia, O., Weber, I., Frangakis, A.S., Nicastro, D., Gerisch, G., Baumeister, W., 2002. Macromolecular architecture in eukaryotic cells visualized by cryoelectron tomography. Science 298, 1209–1213.

Murphy, K., Weiss, Y., 1999. Loopy belief propagation for approximate inference: an empirical study. Proceedings of the 15th Conference on Uncertainty in Artificial Intelligence, pp. 467–475.

Ress, D., Harlow, M.L., Schwarz, M., Marshall, R.M., McMahan, U.J., 1999. Automatic acquisition of fiducial markers and alignment of images in tilt series for electron tomography. J. Electron Microsc. 48, 277–287.

Yedidia, J.S., Freeman, W.T., Weiss, Y., 2005. Constructing free-energy approximations and generalized belief propagation algorithms. IEEE Trans. Inf. Theory 51 (7), 2282–2312.

Yedidia, J.S., Freeman, W.T., Weiss, Y., 2000. Generalized belief propagation. Proceedings of the 14th Conference on Neural Information Processing Systems, pp. 689–695.

Zheng, S.Q., Keszthelyi, B., Branlund, E., Lyle, J.M., Braunfield, M.B., Sedat, J.W., Agard, D.A., 2007. UCSF tomography: an integrated software suite for real-time electron microscopic tomographic data collection, alignment, and reconstruction. J. Struct. Biol. 157, 138–147.

Please cite this article in press as: Amat, F. et al., Markov random field based automatic image alignment ..., J. Struct. Biol. (2007), doi:10.1016/j.jsb.2007.07.007