An accurate, automatic method for markerless alignment of electron tomographic images.

CONFERENCE PAPER · JANUARY 2010

DOI: 10.1109/BIBM.2010.5706597 · Source: DBLP

6 AUTHORS, INCLUDING:

Fa Zhang

Chinese Academy of Sciences

60 PUBLICATIONS **122** CITATIONS

SEE PROFILE

Kai Zhang

Chinese Academy of Sciences

48 PUBLICATIONS **295** CITATIONS

SEE PROFILE

Xiaohua Wan

Chinese Academy of Sciences

11 PUBLICATIONS **25** CITATIONS

SEE PROFILE

Zhiyong Liu

Chinese Academy of Sciences

120 PUBLICATIONS **777** CITATIONS

SEE PROFILE

An accurate, automatic method for markerless alignment of electron tomographic images

Qi Chu*‡, Fa Zhang*, Kai Zhang†‡, Xiaohua Wan*‡, Mingwei Chen*‡, Zhiyong Liu*
*Advanced Computer Research Center,Institute of Computing Technology, CAS, Beijing, China
†National Key Laboratory of Biomacromolecules, CAS, Beijing, China
‡Graduate University of Chinese Academy of Sciences, Beijing, China
Email: chuqi@ict.ac.cn

***Abstract*—Accurate alignment of electron tomographic images without using embedded gold particles as fiducial markers is still a challenge. Here we propose a new markerless alignment method that employs Scale Invariant Feature Transform features (SIFT) as virtual markers. It differs from other types of feature in a way the sufficient and distinctive information it represents. This characteristic makes the following feature matching and tracking steps automatic and more reliable, which allows for estimating alignment parameters accurately. Furthermore, we use Sparse Bundle Adjustment (SPA) with M-estimation to estimate alignment parameters for each image. Experiments show that our method can achieve a reprojection residual less than 0.4 pixel and can approach the same accuracy of marker alignment. Besides, our method can apply to adjusting typical misalignments such as magnitude divergences or in-plane rotation and can detect bad images.**

I. Introduction

Electron tomography (ET) has attracted more and more concerns as it provides the prospect of exhibiting cell structure in molecular detail, e.g. molecular organization of cytoplasm [1], networks of actin filaments in lamelipodia [2]. The first step of ET is to collect projection data. Here we focus on single axis data collection method, that is to tilt around a certain axis in different angles, to project the sample in electron microscopy and finally to record the projection images for computer processing. When being projected, the position of the sample is instable which raises misalignments between projection images that may cause low-quality reconstruction of the sample. Thus the goal of alignment is to retrieve the specific position changes from images and align them by inverse transformations.

Large translational and in-plane rotation changes can be calculated by coarse alignment, using techniques such as cross-correlation [3], common lines [4], or polar cross-correlation [5]. These methods rely on finding the peak location in the Fourier space. The values of the neighborhood of the peak differ little with that of the peak which indicates the images are influenced by high noise and the translational or in-plane rotation estimation is not reliable enough.

Fiducial marker alignment has remained by far the most accurate way to obtain alignment parameters, but unfortunately not every sample can be embedded with particles easily, and the particles may be poorly distributed so that hide the information of sample structure [6].

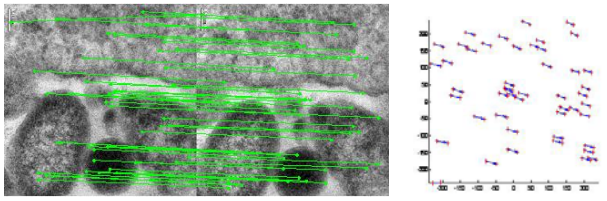
Markerless alignment can be roughly classified into two categories: grey-based alignment and feature-based alignment.

Grey-based alignment attempts to correlate the grey levels of image patches, assuming that they present some similarity. It uses in-process reconstruction-reprojection images compared to original images to correct the position parameters iteratively. This method is well applied to relatively textured images [7]. However it is very computational and better to be speeded up using parallel hardware like clusters or GPUs.

Feature-based alignment is less computational for it uses some points not all points for parameter estimation. Since the principle of feature-based alignment is similar to that of fiducial marker alignment, the differences are: (i) real marker features are substituted by virtual feature points found by artificial intelligence, thus unlike marker features extraction, it is inevitable to generate location errors; (ii) one feature track may only cover a short range of tilt series, and large number of tracks are generated that make the underlying normal equations very large and sparse. Substantial effort has been paid on how to extract and match features reliably. Usually, feature can be determined by informative locations [8] [9] [10] [11], and tracked in corresponding area of adjacent images by normalized cross-correlation iteratively. However, cross-correlation is robust when enduring translational changes and is highly sensitive to changes such as viewpoint change or non-rigid deformations [12] which is just the changes between ET images. This leads to two main errors that highly influence position parameter determination: feature location errors and false matches.

In this paper we propose a new strategy of feature-based alignment using Scale Invariant Feature Transform (SIFT) feature detection [12]. This modern feature detection technique differs from previous methods in that it is represented by a 128-dimensional vector and in the way it matches each other by nearest neighbor algorithm which determines feature matches by searching smallest distance in feature database. SIFT has proven to be invariant to image rotation

978-1-4244-8305-1/10/$26.00 ©2010 IEEE



(a) (b)

Figure 1. Matches of an image pair after bad match elimination. In order to avoid clutter, we only show the first 50 matches between two views. (a) feature matches (b) feature matches in the same coordinate.

and scale and robust toward affine distortion and noise [12]. Furthermore its calculation is efficient so that it has the potential for real-time applications. After using SIFT to extract features of each image and matching and tracking them accordingly, we build up 3D models for each feature track. To solve the large projection model with condition of feature measurement errors, we employ sparse bundle adjustment [13] with M-estimators as robust estimation of position parameters.

II. Method

A. feature extraction

It should be realized that one feature consists of two parts: the distinctive locations, i.e. detector and the surrounding information it represents, i.e. descriptor. Previous research with features extraction has only brought the grey value in the neighborhood of features for cross-correlation, not exploited the abundant information that the neighborhood renders. Our method adopt SIFT descriptor which computes a histogram of local oriented gradients around the feature point and stores the data in a 128-dimensional vector (8 orientation bins for each of the 4×4 location bins). This kind of descriptor is very distinctive as it captures a substantial amount of information about the spatial intensity patterns, and being robust to noise or geometric deformations [12].SIFT is widely used in computer vision field such as object recognition and video tracking. Recently it has been applied to analyzing 3D magnetic resonance images of the human brain to detect Alzheimer's disease [14], whose results show a high ratio recognition of the disease in this kind of low signal-to-noise ratio images comparable to ET images.

One 512×512 pixel image can give rise to about 2000 SIFT features distributed around areas rich in detail. The number of features is sufficient for generating large amount of feature tracks covering the whole series. The location coordinate and the 128-dimensional vector of each feature are then stored in a database.

B. feature matching

Finding feature correspondences between two views can be done by finding minimum Euclidean distance between feature vectors in database. Only those correspondences whose minimum distances are less than a certain threshold can remain. This operation is executed between every image pair. We only record image pairs which have more than 16 correspondences.

To exclude false matches between image pairs, we use epipolar constraint. Those matches which generate large residuals multiplied with the constraint are pruned and those generate small residuals are corrected.

C. feature tracking

Feature tracks are formed by threading the feature matches from the last step. The key skill is the computer searching skill. We check the set of image pairs one by one. For each feature in one image, find its correspondence in neighborhood images iteratively. One feature which has at least one correspondence forms a feature track and one feature track is assumed to represent a projection sequence of a certain virtual point in 3D space.

D. parameter estimation

The relation between the virtual 3D point and its projection in 2D image can be formalized as (1) [3],which is widely used in markerless alignment:

$$A_j = s_j R_z(\alpha_j) P R_y(\theta_j) R_x(\varphi) \tag{1}$$

Where s_j is the scaling factor, $R_z(\alpha_j)$ is the in-plane rotation, P is orthographic projection matrix, $R_y(\theta_j)$ and $R_x(\varphi_j)$ are the rotations around y-axis and x-axis respectively. The rotation parameters can be described using quaternion form [15] to allow for big parameter changes although it is unnecessary in this case as true values are near the initial values.

In this kind of parameter estimation problem which is also called Structure and Motion(SAM) problem in stereo vision, we treat structure and motion parameters uniformly and estimate them using well known least-squares method:

$$\{s_j, \alpha_j, \theta_j, \varphi, r_i\} = \underset{s_j,\alpha_j,\theta_j,\varphi,r_i}{\operatorname{argmin}} \sum_{i,j} |m_{ij} - \hat{m_{ij}}|^2 \tag{2}$$

Classical methods for optimizing the parameters include conjugate gradients, linear regression, Newton-like methods and adaptive methods such as trust region method. Those methods are prune to outliers. Thus we employ sparse bundle adjustment [13] that jointly optimize 3D feature coordinates and projection parameters and take advantage of the sparsity of the gradient to reduce computation and memory storage. At each round, we get a set of parameters and residual $|m_{ij} - \hat{m_{ij}}|$. Then we perform M-estimators to exclude m_{ij} which has large residual.

M-estimation is formalized by replacing the minimization objective with objective function:

Table I
TECHNICAL DETAILS

	centriole	mitochondrion	caveola
Voltage(kV)	300	200	200
Full image size	512×512	1024×1024	512×512
No. Views	61	112	119
Pixel Width(nm)	2.02	0.8	1.6
No. Features(/image)	1312	20963	3030
No. Chains	9270	18869	4155
Aver. Chain Len.	3	2	2
Max. Chain Len.	57	29	22
No. Rounds	3	4	3
No. Outlier Chains	769	5198	287
No. Views Excluded	0	0	3
Reprojection Res.	0.242	0.381	0.311

$$E = \sum_j \rho(r_j(m_j, P); \sigma) \tag{3}$$

Here, r_j is residual, P represents the model, σ is the error threshold we choose. Rounds continue until no m_{ij} can be excluded.

III. EXPERIMENTS AND RESULTS

We present three real series of data for analysis. The first series is downloaded from IMOD homepage [16], which is experimental data for fiducial marker alignment. We use it as a reference to our markerless alignment. The other two are recorded by National Key Laboratory of Biomacromolecules. Coarse alignment is carried out first in default, which includes large-shifts correction using cross-correlation and y-axis correction using Random transform of power spectrum. Thus all the comparisons below are between result of coarse alignment and that of our method. We use the popular weighted back projection(WBP) [17] for reconstruction.

The first experiment is centriole sample digitized by Gatan Camera on TF30 with voltage 300kV. Markerless alignment details are in Table 1. The table includes basic information of the three data sets, and feature tracking and parameter estimation information accordingly. The mean reprojection residual of the first series is 0.242 pixel compared to 0.321 pixel in marker alignment as we set up a threshold in M-estimator that gross errors will be eliminated in each round. The small residual means that the estimated parameters are very close to the measured parameters, just a reminder, not the true ones. By eye-checking Fig.2, when only apply coarse alignment, we can see dispersion of fiducial markers which is caused by shift misalignment, whereas after marker alignment or our method, dispersion does not exist. Besides, the same cross sections of reconstruction show that tubulin dimers of microtubules are clearly rendered after applying marker method or our method other than coarse alignment.

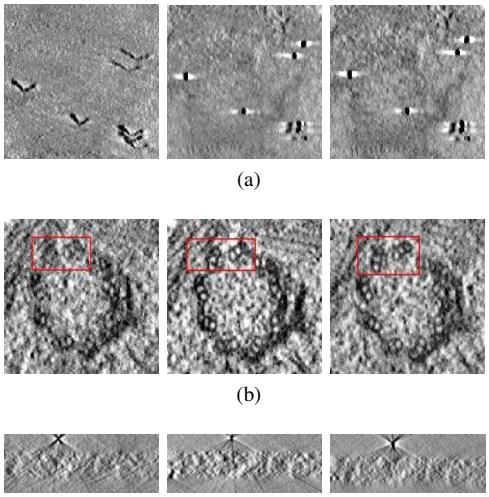


Figure 2. From left to right are the reconstruction section after coarse alignment, marker alignment and our robust markerless alignment. (a) is top cross section, from which we can see the wing of gold particles in coarse alignment caused by shift misalignment. Also longitudinal section (c) illustrates the same effect. (b) is an area of center cross section harboring useful information. Tubulin dimers of microtubules can be seen clearly after applying marker alignment and our method.

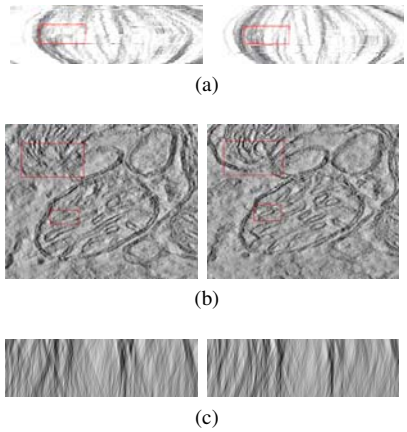
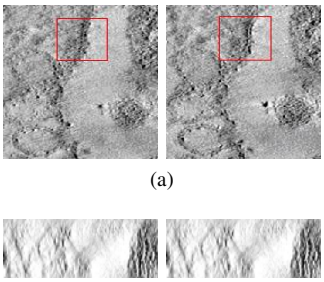


Figure 3. Comparison between coarse alignment (left) and our method (right) of mitochondrion. (a) is the longitudinal section of coarse aligned images and aligned images using our method respectively. Inconsistent sections means scale variations which are corrected by our method shown on the right. (b) and (c) represent cross and longitudinal sections of reconstruction result.

However our method is fully automatic while marker alignment needs to select marker seeds by hand first.

The second and third data sets are both colleted by FEI company's production– Tecnai 20 [18]. The second test sample is mitochondrion of mice hepatic cell. With a glance of original data, we can see variations in scale about $\pm 2\%$ and in in-plane rotation changes which are about $\pm 2.5^0$. The scale variations are rendered in figure 3a with discontinuation of sections. After applying our method, the discontinuation is smoothed meaning that the scale and rotation changes have been adjusted well. The reconstruction



(a)

(b)

Figure 4. Comparison between coarse alignment(left) and our method(right) of caveola. (a) is cross section and (b) is longitudinal section. From our method we can see legible phospholipid bilayer.

result shows that the ridge of mitochondrion is more clear than only using coarse alignment.

The third test sample focuses on caveola of porcine aorta endothelial cell(PAE cell) [19]. This series has a large noise level. After three rounds of computation, three views that have been blurred too much to have information more than noise have been detected by our method and are to be excluded before reconstruction. Reconstruction result shows that after using our method, the structure–phospholipid bilayer is legibly unveiled.

IV. Conclusion and Future work

Alignment quality is crucial for reconstruction resolution of ET images. How to retrieve the alignment parameters from images without using fiducial markers is still a challenge. In this paper we propose a new accurate markerless alignment method which introduces SIFT features as a substitution of fiducial markers. It has good attributes such as invariant to image scale and rotation, affine distortion, and noise [12]. The matching strategy of our method contrasts with other correlation-based strategies in that it finds the smallest distance of a selected feature from a database of other features, and more importantly SIFT is distinctive enough to enable the correct match. After matching and tracking features throughout the whole series, we use SPA with M-estimation to obtain position parameters. Compared to other accurate alignment methods like [8], the reprojection residual of our method is much smaller which means our parameter estimation is more accurate with respect to the same kind of measured location values.

As shown from the experiments, our method is well applied to varied misalignment including random shifts, scale, or in-plane variations and can detect bad images in a series; our method is obviously more accurate compared to coarse alignment and can achieve the accuracy of marker alignment.

As modern electron microscopy like Gatan with automatic data collection equipment can generate ET data in the magnitude of terabyte in one day, fast or even real-time alignment is unavoidable. Thus our future work will rest on reducing the complexity of our method in each step in order to meet real-time requirements.

References

[1] W. Baumeister, *Current Opinion in Structural Biology*, vol. 12, pp. 679–684, 2002.

[2] E. Urban *et al.*, *Nature Cell Biology*, vol. 12, no. 5, pp. 429–436, 2010.

[3] J. Frank and B. McEwen, in *Electron Tomography:Three-Dimensional Imaging with the Transmission Electron Microscope.*, J. Frank, Ed. Plenum Press, 1992, pp. 205–213.

[4] Y. Liu *et al.*, *Ultramicroscopy*, vol. 58, no. 3-4, pp. 393–402, 1995.

[5] J. Frank, in *Computer Processing of Electron Microscope Images*, P. W. Hawkes, Ed. Springer-Verlag, 1980, pp. 187–222.

[6] M. David, N., in *Electron Tomography:Methods for Three-Dimensional Visualization of Structures in the Cell.*, J. Frank, Ed. NY: Springer, 2006.

[7] H. Winkler and K. Taylor, *Ultramicroscopy*, vol. 106, no. 3, pp. 240–254, 2006.

[8] J. H. Brandt, S. and P. Engelhardt, *Journal of Structural Biology*, vol. 136, no. 3, pp. 201–213, 2001.

[9] D. Castano-Diez *et al.*, *Journal of Structural Biology*, vol. 159, no. 3, pp. 413–423, 2007.

[10] S. Phan *et al.*, *Proceedings of the 2009 WRI World Congress on Computer Science and Information Engineering*, pp. 604–612, 2009.

[11] C. Sorzano *et al.*, *Bmc Bioinformatics*, vol. 10, 2009.

[12] D. Lowe, *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[13] M. Lourakis and A. Argyros, *Acm Transactions on Mathematical Software*, vol. 36, no. 1, 2009.

[14] M. Toews *et al.*, *Neuroimage*, vol. 49, no. 3, pp. 2318–2327, 2010.

[15] J. Kuipers, in *Quaternions and Rotation Sequences:A Primer with Applications to Orbits,Aerospace and Virtual Reality.*

[16] Imod homepage. [Online]. Available: http://bio3d.colorado.edu/imod/

[17] G. Herman, *Image Reconstruction from Projections.*, Academic Press, 1980.

[18] Tecnai 20 operating manual. [Online]. Available: http://sbio.uct.ac.za/Webemu/documentation_technical/equipment/equipment_manuals/fei-f20_f30/Tecnai_20_Operating_Manual.pdf

[19] S. Sun *et al.*, *Journal of Microscopy*, vol. 36, no. 6, pp. 729–735, 2009.