MVA2000 IAPR Workshop on Machine Vision Applications, Nov. 28–30, 2000, The University of Tokyo, Japan

8—16

A Fully Automatic Alignment of Electron Tomography Images Without Fiducial Markers

Sami Brandt*and Jukka Heikkonen†
Laboratory of Computational Engineering
Helsinki University of Technology

Abstract

In structural biology, electron tomography is used in reconstructing three-dimensional objects such as macromolecules, viruses, and cellural organelles to learn their structure and properties. In order to successfully perform the three-dimensional reconstruction from a series of transparent two-dimensional images, the images have to be aligned or registered. In this paper, we propose a multi-phase method where the registration process is fully automated without the need for fiducial markers such as gold particles. Our experiments show promising results.

1 Introduction

Electron tomography means reconstructing the interior of an object from its electron microscope images. In order to perform the 3D reconstruction successfully, the motion between the successive images must be solved, i.e., the images have to be aligned or registered. In this paper we use a set of two-dimensional electron microscope images of a three-dimensional object which are obtained by tilting the microscope around one axis during the imaging.

Traditionally the alignment is solved either by manually showing the corresponding markers from the set of images [7] or automatically using simple correlation between the images on several rotations and scale [5]. The manual approach has the disadvantage of being time consuming for the user because, at minimum, the number of shown markers tends to be hundreds. The previous correlation-based automatic methods are, however, much more inaccurate because the transformations are computed between the successive images of the tilt series and therefore the errors cumulate along the images.

We propose a method where the registration is automated using the state-of-the-art computer vision methods but, as in our previous work [1], no fiducial markers are needed. The solution to the problem lies on searching for point correspondences which are tracked along the images which makes it possible to optimize the motion parameters for the whole image set at time. Therefore, our method has shown promising results to meet the difficulties of the present registration methods in electron tomography.

The registration problem is divided to several subproblems: (1) finding initial matches from successive images, (2) estimating the epipolar geometry between images, (3) refining and predicting the probable matches using the epipolar geometry with the disparity information, (4) matching and tracking the refined matches through the tilt series, and (5) optimizing the transformation parameters for the image set.

2 Determining the Initial Correspondence

In establishing correspondence between images, we will use epipolar geometry as the matching constraint. To estimate the epipolar geometry we, however, need some initial matching points between images. To solve the initial correspondence we use the Harris corner detector [6] followed with the correlation and relaxation techniques described in [8]. Because the heuristical correlation and relaxation process may also give lots of false matches we have proposed another post-processing step which reduces the number of false matches [1]. It is based on the observation that the flow, which is computed by taking the difference between the match coordinates in two images, must be consistent with the flow of the neighboring matching points. Therefore, matches associated with flow vectors which are not supported by others are discarded. For details, see [1].

3 Epipolar Geometry Estimation

In the second stage, the epipolar geometry between successive images is estimated. Epipolar ge-

*Address: P.O. Box 9400, FIN-02015 HUT, Finland. E-mail: Sami.Brandt@hut.fi
†Address: P.O. Box 9400, FIN-02015 HUT, Finland. E-mail: Jukka.Heikkonen@hut.fi

ometry provides the only geometrical constraint between two images taken from the same scene. Once the epipolar geometry is known and given one point in one image, one can immediately tell the corresponding line in the other image where the corresponding point lies. The constraint is represented by a 3×3 matrix called fundamental matrix (see e.g. [8]).

Because an electron microscope can be accurately approximated with an affine camera model, we use our novel Bayesian approach for the fundamental matrix estimation [2]. The method, instead of classifying the matches to relevant and false, weights the matches by their a posteriori probability to be relevant. More accurately, the estimation is performed as follows,

At first an initial estimate is computed using the LMedS method for the affine F-matrix estimation [9]. With this estimate the residual may be computed from

$$\epsilon_i = \frac{\mathbf{u}_i^{\mathrm{T}}\mathbf{f} + f_{33}}{\sqrt{\mathbf{f}^{\mathrm{T}}\mathbf{f}}}, \tag{1}$$

where $\mathbf{f} = (f_{13} \;\; f_{23} \;\; f_{31} \;\; f_{32})^{\mathrm{T}}$ and $\mathbf{u}_k = (\mathbf{m}_i{}^{\mathrm{T}} \;\; \mathbf{m'}_j{}^{\mathrm{T}})^{\mathrm{T}}$. We have proposed [2] to model the residual by a Gaussian mixture model,

$$p(\epsilon) = P_{\mathrm{r}}\, p(\epsilon|S_{\mathrm{r}}) + P_{\mathrm{f}}\, p(\epsilon|S_{\mathrm{f}}), \tag{2}$$

where the two Gaussian components $p(\epsilon|S_{\mathrm{r}})$ and $p(\epsilon|S_{\mathrm{f}})$ correspond to the residuals of the relevant and false matches, respectively. The parameters are solved by maximizing the likelihood function

$$L = \prod_{i=1}^{N} p(\epsilon_i|P_{\mathrm{r}}, P_{\mathrm{f}}, \mu_{\mathrm{r}}, \mu_{\mathrm{f}}, \sigma_{\mathrm{r}}, \sigma_{\mathrm{f}}) \tag{3}$$

with respect to its parameters. Now, the a posteriori probability $P(S_{\mathrm{r}}|\epsilon)$ is given by the Bayes rule

$$P(S_{\mathrm{r}}|\epsilon) = \frac{P_{\mathrm{r}}\, p(\epsilon|S_{\mathrm{r}})}{P_{\mathrm{r}}\, p(\epsilon|S_{\mathrm{r}}) + P_{\mathrm{f}}\, p(\epsilon|S_{\mathrm{f}})}, \tag{4}$$

which is used in computing a new estimate for the fundamental matrix. This is performed by weighting the new residual by the a posteriori probabilities to be relevant given the old residuals which is equivalent to solving the following eigen equation and taking the eigen vector corresponding to the smallest eigen value:

$$\mathbf{W}\mathbf{f} - \lambda\mathbf{f} = 0, \tag{5}$$

where $\mathbf{W} = \sum_{i=1}^{N} P(S_{\mathrm{r}}|\epsilon_i)(\mathbf{u}_i - \mathbf{u}_0)(\mathbf{u}_i - \mathbf{u}_0)^{\mathrm{T}}$, $\mathbf{u}_0 = \frac{1}{N}\sum_{i=1}^{N} \mathbf{u}_i$. The parameter f_{33} is obtained from

$$f_{33} = -\mathbf{u}_0^{\mathrm{T}}\mathbf{f}. \tag{6}$$

Another advantage of the Bayesian method is that we also obtain a direct estimate for the fundamental matrix covariance from [3]

$$\mathbf{C}_{\mathbf{f}} \simeq \sigma^2\mathbf{Q}\Big(\sum_i P(S_{\mathrm{r}}|\epsilon_i)^2(\mathbf{u}_i - \mathbf{u}_0)(\mathbf{u}_i - \mathbf{u}_0)^{\mathrm{T}}\Big)\mathbf{Q}^{\mathrm{T}}, \tag{7}$$

where

$$\mathbf{Q} = -\sum_{k=2}^{4} \frac{\mathbf{q}_k\mathbf{q}_k^{\mathrm{T}}}{\lambda_k}, \tag{8}$$

whereas λ_k and $\mathbf{q}_k$ are the k^{th} largest eigenvalue and its associated eigenvector of $\mathbf{W}$, respectively.

The covariance estimate is here used in the next phase as it contains information for computing the confidence intervals of the epipolar lines and even some disparity information of the images.

4 Finding candidates

In general, if the epipolar geometry was not used in the matching process we should consider every corner pair as a possible match. We may, however, use the epipolar geometry to reduce the number of possible matches substantially. If the F-matrix was known accurately, we might consider points only at the epipolar lines, but because there is always errors in the F-matrix estimation we search for the epipolar bands or the error bounds of the epipolar lines which are determined by the F-matrix covariance matrix estimate (7) [3].

The epipolar band is usually a hyperbola on both sides of the epipolar line. The covariance matrix contains also some disparity information of the image since the bounds are narrowest at the location where the match most probably lies [4]. Sometimes the error bounds may also be far too pessimistic, so here we search only the intersection of the epipolar band and the rectangular area centered at the most probable point with width and height 250 and 50 pixels, respectively. The corners that are found in this region in the second image are set as candidate matches for the corner in the first image. Usually there are no more than a couple of possible candidates thus the use of the epipolar geometry reduces the possibility of a false match greatly.

5 Graph Matching and Tracking

The neighborhood graphs are next formed for the corners similarly as explained in [1]. The graphs are here associated with the coordinates of the reference corner and its k nearest neighbors.

The corresponding graphs in the second image are sought for each corner using the above solved candidate set. In order to perform this we should take the following things in the account: Firstly, to maximize the separability of the classifier, the graph

matching procedure should neither be scale nor rotation invariant because we have the a priori knowledge that successive images do not differ much from each other. However, we should have robustness for both rotation and scale. Secondly, the images may, however, have a small shift in rotation due to the manual scanning from the film.

Therefore, to remove the effect of the shift in rotation, we first transform the graphs such that the epipolar lines are horizontal. As the affine camera model is used this is achieved by a simple rotation of the graph where the angle can directly be determined by the fundamental matrix [8].

After the coordinate transform, for each graph, a discrete 2D neighborhood uncertainty or impulse map is formed: In the reference corner centered coordinate system we set the values one for the locations of the neighboring corners and zero anywhere else. Between the images, the location uncertainties are modeled by a symmetric Gaussian distribution with variance σ thus the maps are convolved with a two-dimensional Gaussian kernel.

The purpose of the above maps is to code the location information of the corners in a robust way where the neighboring information is preserved. The comparison of the graphs is made by computing the normalized correlation between the uncertainty maps of the matching candidates. Now, the maximum correlation scores exceeding 50% are taken as correspondences. In addition, we use values 5 and 10 for k and combine the results since more matches can be found in this way compared to a single value of k.

The above computations are finally made for all consecutive image pairs in a row and the results are saved in an appropriate data structure. The corner tracking is thereafter rather straightforward to implement since all that must be found anymore is the individual corner coordinates in each image and the image numbers where the chain starts and ends.

6 Parameter Optimization

The final stage in solving the alignment is the optimization of the parameters which are needed in transforming each image to a common coordinate plane. As the imaging operation model we use the one defined by Lawrence [7] with slight modifications. The j^{th} corner coordinates $\mathbf{m}_j^i = (x_j^i \; y_j^i)^{\mathrm{T}}$ in the i image are related to the corresponding 3D coordinates $\mathbf{x}$ as follows:

$$\mathbf{m}_j^i = s^i \mathbf{R}_\alpha^i \mathbf{P} \mathbf{R}_\beta^i \mathbf{x}_j + \mathbf{t}^i, \tag{9}$$

where s^i is a scaling factor, $\mathbf{R}_\alpha^i$ is a 2×2 rotation matrix associated with the angle α, $\mathbf{R}_\beta^i$ is 3×3 rotation matrix describing the tilting operation around y-axis, $\mathbf{t}^i$ is a translation vector for image i, and $\mathbf{P}$

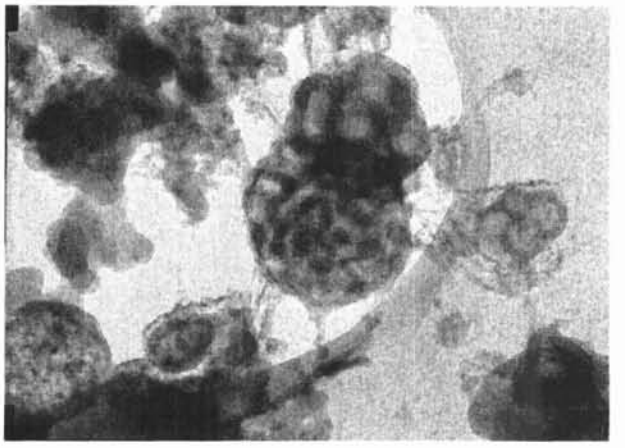

Figure 1: An electron microscope image of a mitochondrion.

is an orthographic projection matrix

$$\mathbf{P} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}. \tag{10}$$

Optimal estimates for the unknown transformation parameters are obtained by minimizing the sum of residuals squared, which is equivalent the following cost function being minimized

$$Q(x) = \sum_i \sum_j \left(\hat{\mathbf{m}}_j^i - \mathbf{m}_j^i\right)^{\mathrm{T}} \left(\hat{\mathbf{m}}_j^i - \mathbf{m}_j^i\right) \delta_{ij}, \tag{11}$$

where $\hat{\mathbf{m}}_j^i$ is the measured corner coordinate vector, and δ_{ij} is the Kronecker delta product indicating whether the j^{th} indexed corner chain is found in image i. The unknown the parameters are finally optimized by minimizing the cost function using standard optimization tools.

7 Experiments

In the first experiments we have used a tilt series of a mitochondrion of which one image is displayed in Figure 1. The image series consist of 41 images tilting the specimen by three degree increments from $60°$ to $60°$ where one image size is about 3200×2300 pixels. Some results are illustrated in Figure 2 where few trajectories of the found feature points are shown after the transformations are made to the images. Ideally, the trajectories should be horizontal lines which is quite well fulfilled here.

As a remark, it is known that the localization of the Harris corner detector is not very accurate. In addition, corner point localization is a difficult problem for this kind of a natural object. In order to get better results, a more sophisticated corner or some other feature detector should therefore be experimented. In addition, as it can be seen form

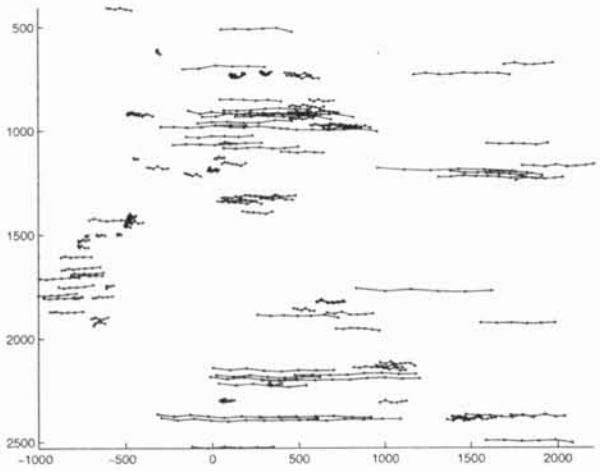

Figure 2: Some feature point trajectories in the tilt series after the registration. One feature point location is marked with points while the locations of the same feature point in different images are connected with lines. Here only the chains longer than 5 are shown for the clarity reasons.

the Figure 2, the chains are relatively short which also means that the corner features are relatively sensitive to the object transformation thus difficult to track. This is naturally shown in the accuracy of the final result.

8 Conclusions

The purpose of this work was to bring the present knowledge of computer vision for automating the alignment process of electron tomography images. We have succesfully found a solution to the problem including that we have developed a new robust method for the affine fundamental matrix and its covariance estimation, developed a certain graph matching technique for finding correspondences reliably, and solved some difficulties in coping with the large number of parameters in the transformation optimization. The results obtained so far are promising and they indicate that electron tomography images can be aligned using corner features in the way proposed. In future, we are going to develop the matching stage from a more general point of view in order to achieve better accuracy and to make the method applicable in the general matching problem of multiple views.

Acknowledgements

We would like to thank Dr. Peter Engelhardt from Department of Virology, Helsinki University for the test images and co-operation.

This work has been supported by the Graduate School in Electronics, Telecommunications, and Automation (GETA) and Academy of Finland, Research Centre for Computational Science and Engineering, project. no. 44897 (Finnish Centre of Excellence Programme 2000–2005).

References

[1] Sami Brandt and Jukka Heikkonen. Automatic alignment of electron tomography images using markers. In David P. Casasent, editor, *Intelligent Robots and Computer Vision XIX: Algorithms, Techniques, and Active Vision*, volume 4197 of *SPIE Proceedings Series*, Boston, USA, November 2000. To appear.

[2] Sami Brandt and Jukka Heikkonen. A new robust Bayesian method for the affine F-matrix estimation. In *5th Fall Workshop Vision, Modeling, and Visualization 2000*, Saarbrücken, Germany, 2000. To appear.

[3] Sami Brandt and Jukka Heikkonen. A robust Bayesian method for the affine fundamental matrix and its uncertainty estimation. Submitted to *Computer Vision and Image Understanding*, 2000.

[4] Gabriella Csurka, Cyril Zeller, Zhengyou Zhang, and Olivier D. Faugeras. Characterizing the uncertainty of the fundamental matrix. *Computer Vision and Image Understanding*, 68(1):18–36, October 1997.

[5] Joachim Frank and Bruce F. McEwen. Alignment by cross-correlation. In Joachim Frank, editor, *Electron Tomography; Three Dimensional Imaging with the Transmission Electron Microscope*, volume 8, pages 205–213. Plenum Press, New York, USA, 1992.

[6] C. Harris and M. Stephens. A combined corner and edge detector. In *4th Alvey Vision Conference*, pages 147–151, 1988.

[7] Michael C. Lawrence. Least-squares method of alignment. In Joachim Frank, editor, *Electron Tomography; Three Dimensional Imaging with the Transmission Electron Microscope*, chapter 8, pages 197–204. Plenum Press, New York, USA, 1992.

[8] Gang Xu and Zhengyou Zhang. *Epipolar Geometry in Stereo, Motion and Object Recognition*, volume 6 of *Computational Imaging and Vision*. Kluwer Academic Publishers, 1996.

[9] Zhengyou Zhang. WWW home page. <URL: http://www- sop.inria.fr/robotvis/personnel/zzhang/zzhang-eng.html>, 1999.