On the Alignment of Transmission Electron Microscope Images Without Fiducial Markers

Sami Brandt and Jukka Heikkonen
Helsinki University of Technology
Laboratory of Computational Engineering
P.O. Box 9400, FIN-02015 HUT, Finland
Sami.Brandt@hut.fi, Jukka.Heikkonen@hut.fi

Peter Engelhardt
University of Helsinki, Haartman Institute,
Department of Virology, P.O. Box 21,
FIN-00014 University of Helsinki, Finland
Peter.Engelhardt@helsinki.fi

Abstract
Accurate image alignment is needed for computing three-dimensional reconstructions from transmission electron microscope tilt series. So far, the best alignment results have been obtained by using colloidal gold beads as fiducial markers. If their use has not been possible for some reason, the only option has been the automatic 2D-cross-correlation-based registration methods. However, since the actual motion is three dimensional this approach is inaccurate and inappropriate for the whole problem. Conversely, we propose a novel method that uses the actual 3D motion model without any fiducial markers in the images. The method is based on matching and tracking corner features by first solving the underlying geometrical constraint of consecutive images in the tilt series. The results show that our method is competitive with the gold marker alignment in the level of accuracy and hence opens the way for new opportunities in the analysis of electron tomography reconstructions.

1. Introduction
Aligning transmission electron microscope (TEM) images in a most accurate way for electron tomography is important. The resolution of the three-dimensional reconstruction is directly affected by the success of the registration as inaccuracies lead to increasing reconstruction noise occurring, e.g., as blurring and stretches of the reconstruction.

Accurate alignment has been previously achieved by using conventional colloidal gold beads as fiducial markers. The accuracy is due to spherical-shaped gold markers that can be well localized. In fact, until recently these approaches have been manual where the user must have shown the locations of the markers in the images. A clear disadvantage has been the grinding need for user interaction that practically limits the number of markers used substantially. Some recent approaches [6, 10, 3] have consequently considered automating the gold marker picking process.

Sometimes it is not, however, possible to have gold markers in the images. Gold markers also interfere with computation of the reconstruction so it would occasionally be necessary to solve the image alignment without markers. As a matter of fact, electron tomography literature knows automatic alignment by cross-correlation [7, 5] where no fiducial markers are needed. However, that approach is only based on the 2D cross-correlation that is capable of modeling 2D plane motion—not the 3D motion of the object. Since the implicit motion model of this approach is simply insufficient for the whole alignment problem, the result is not optimal at all and the alignment errors accumulate along the image series.

In this paper, we introduce a framework for registering TEM images that for the first time provides the automatic alignment using the actual 3D motion model without any fiducial markers. The approach is based on first estimating the epipolar geometries of the consecutive image pairs after which the found feature points are tracked using a wavelet-based multi-resolution approach. The maximum likelihood estimates for the transformation parameters are finally obtained by solving a non-linear optimization problem. This paper hence reports how a carefully considered combination of computer vision techniques can efficiently solve the alignment problem where no user-interaction is needed.

2. Methods
Our approach starts with corner detection (Section 2.1) and thereafter estimating epipolar geometries between consecutive image pairs in the tilt series (Section 2.2). In matching we use a wavelet-based multi-resolution technique (Section 2.3) after which motion parameter optimization follows (Section 2.4).
Figure 1. Examples of Harris corners found in the four reference images used in the experiments, see Fig. 2. The mark ‘x’ indicates a found corner and ‘+’ a corner for which a matching point has been found in next image. Sizes of the windows are 125 × 125 pixels.

2.1. Feature Point Extraction

In motion estimation, the usual approach is to establish correspondences between feature points in the images and use some explicit motion model whose parameters are to be estimated on the basis of the measurements. In general, good interest points are points that can be accurately localized and whose locations are stable and invariant to the changes in the imaging conditions. Gold markers would fulfill these requirements well [3] but, as we now assume, their use is not always possible. We therefore track corner features obtained (see Fig. 1) by the improved Harris detector [11] due to its good repeatability and localization accuracy. Now, the main part of the alignment problem is establishing the correspondences between the found corners.

2.2. Epipolar Geometry Estimation

The only geometrical constraint that involves two views is called the epipolar constraint, see e.g. [8]. Known epipolar geometry facilitates the matching problem substantially since, given a point in the one image, the correspondence will be found on the epipolar line in the other image. Epipolar geometry is represented by a $3 \times 3$ matrix called the fundamental matrix. However, initially the epipolar geometry and thus the fundamental matrix is unknown and point correspondences are needed for its estimation.

The initial corner point matchings between the subsequent view pairs in the tilt series are established using the correlation and relaxation techniques proposed in [12]. Because these techniques rely on heuristics, the initial correspondences also consist of mismatches and we must use some robust techniques to estimate the fundamental matrix. For this purpose we use our method [2] which does not only give a maximum likelihood estimate for the affine fundamental matrix regardless of the mismatches but also an estimate for the fundamental matrix covariance. The uncertainty information will also be utilized as explained below.

2.3. Multi-Resolution Matching

A classical way in matching interest points is to measure correlation between small windows surrounding the points. Normalized correlation is also a good measure for intensity similarity for electron microscope images. However, even if the estimated epipolar geometry is used as the constraint it is namely possible that a false, matching window may also be found on the epipolar line. To reduce the amount of false matches one could select larger correlation windows—however, the amount of badly localized, nearly correct mismatches would increase.

To cope with the problems above, we have proposed a wavelet-based multi-resolution method for matching [1]. In brief, first a wavelet decomposition for images is computed when we obtain a pyramid-like multi-resolution representation of the images. Starting from the lowest resolution image a constant-sized correlation window is selected around a feature point. All strong correlation maxima in the second image which lie inside the error bounds of the corresponding epipolar lines are recursively investigated by zooming to the higher resolution and searching for further maxima inside the lower resolution window. If for some point correlation is statistically significant on each resolution levels and a corner point is found sufficiently near from the found location we interpret having found a correspondence. If we still have multiple candidates for correspondence we use the covariance of the F-matrix to select the most probable. Technical details can be found from [1, 4].

The proposed procedure uses the neighborhood information of the interest points in a robust way without any retardation in the localization accuracy. As a matter of fact, our experiments have shown that the probability for obtaining mismatches practically vanishes. Being pedant, there however remains a tiny probability for obtaining mismatches. Would this be a problem, as a post-processing step one

\[1\] The affine camera model is practically ideal for TEM.
could additionally consider three view correspondences, estimate the affine trifocal tensor and reject the mismatches on the basis of the point transfer error [8]. Nevertheless, in this work we did not see this necessary since there seemed to be no mismatches in the series used in the experiments.

2.4. Parameter Optimization

Finally, the transformation parameter optimization stage completes the method. The cost function is minimized

\[
Q(x) = \sum_i \sum_j (\hat{m}_j^i - m_j^i)^T (\hat{m}_j^i - m_j^i) \delta_{ij},
\]

where \(m_j^i\) is the measured point coordinate vector, \(\delta_{ij}\) is the Kronecker delta product indicating whether the \(j^{th}\) indexed point chain is found in image \(i\), and the motion model is

\[
\hat{m}_j^i = s^i R_j^i P R_j^i x_j + t^i,
\]

where \(s^i\) is a scaling factor, \(x_j\) is the 3D coordinate vector of the chain \(j\), \(R_j^i\) is a \(2 \times 2\) rotation matrix associated with the angle \(\alpha^i\), \(R_j^i\) is a \(3 \times 3\) rotation matrix describing the tilting operation around the \(y\)-axis, \(t^i\) is a translation vector for image \(i\), and \(P\) is an orthographic projection matrix. The optimization with respect to the parameters \(\alpha^i, \beta^i, s^i\), and \(x_j\) is performed using the classical Gauss–Newton algorithm.

3. Experiments

We tested four different tilt series. The first series is a whole-mounted critical-point-dried chromosome that also contains gold markers that provides us a ground truth to which we may compare our new method though we do not use the markers. The second series is an example from a situation where the markers have been sprinkled on the preparation but where it is not possible to perform a manual alignment due to poor visibility of markers. To demonstrate the viability of the proposed method we have reconstructed this set and in addition two other tilt series that are absolutely free of markers. As the reconstruction method we have used the well known maximum entropy method (MEM) [9].

The whole-mounted critical-point-dried chromosome tilt series (Fig. 2a) originally consists of 41 images taken by 3 degree increments from \(-60^\circ\) to \(60^\circ\) with 100 kV voltage and 48 000x magnification. To evaluate the alignment we compute the standard deviation \(\sigma\) of the difference between the measured corner locations and predicted by the estimated model (2). We may compare this with the corresponding deviation of the gold marker coordinates computed by using the gold marker alignment.

In the first series, by using gold markers we obtained 1.3 pixels for the marker deviation. The alignment without the use of the markers resulted \(\sigma = 2.5\) pixels indicating that, with this image set, the noise level in the gold marker coordinates were about half of the corner points. Nevertheless, this is no surprise because gold markers can be very accurately localized.

The second test set is a typical example where the proposed alignment method could be applied. The object in the Fig. 2b represent a microvillus (taken by 8000x magnification, 120 kV voltage). For this set the corner point deviation was found to be 0.8 pixels after the optimization. This is a very good result considering the difficulty of the image series. The stereo reconstruction of the top part is shown in Fig. 3b.

In Fig. 2c we have a whole-mounted critical-point-dried mitochondrion without markers. The series was taken at 50 000x magnification and 120 kV voltage. The alignment resulted \(\sigma_c = 3.4\) pixels for the deviation. The large deviation is due mostly to some serious dynamic deformation of the object which can also be clearly seen from the tilt series. However, the stereo image pair of the reconstructed mitochondrion in Fig. 3c shows that the proposed alignment method has worked out successfully.

The last tilt series is of a slice of another mitochondrion shown in Fig. 2d. The series of 36 images was taken at

Figure 2. Example images from the tilt series. (a) Chromosome scaffold, (b) freely dangling microvillus, (c) whole-mounted mitochondrion (d) section of a mitochondrion.
Figure 3. Stereo reconstructions. (a) Top part of the microvillus, (b) mitochondrion.

Table 1. Standard deviation of the localization error after the optimization. There were typically hundreds of chains longer than 4.

<table>
<thead>
<tr>
<th>Method</th>
<th>Chr.</th>
<th>Microv.</th>
<th>Mitoch.</th>
<th>Section</th>
</tr>
</thead>
<tbody>
<tr>
<td>With Gold</td>
<td>1.3</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Without Gold</td>
<td>2.5</td>
<td>0.80</td>
<td>3.4</td>
<td>1.3</td>
</tr>
</tbody>
</table>

40 000 × magnification and 80 kV voltage. Now, the alignment resulted 1.3 for the deviation of the corner points. The result is surprisingly accurate even though the noise level in the images is considerably large. The results are summarized in Table 1.

4. Conclusions

We have proposed a novel way for aligning transmission electron microscope images automatically where no fiducial markers are needed. The solution lies on first solving the underlying epipolar constraint between consecutive images after which corner features are tracked using wavelet-based multi-resolution matching approach. The experimental results reveal that the proposed alignment is close to the accuracy level of the alignment where fiducial gold markers are used. Being just on the prototype level, the future development should guarantee even better reliability and accuracy with possible real-time computational efficiency. Therefore it has a promising capability of substituting the previous methods in TEM image alignment for electron tomography. Moreover, it is the only model based alignment method if there are no markers available in the images.

References