Probabilistic Image Registration and Anomaly Detection by Nonlinear Warping

Verena Kaynig[1,2], Bernd Fischer[1], Joachim M. Buhmann[1]
{verena.kaynig, bernd.fischer, jbuhmann}@inf.ethz.ch

[1]Institute of Computational Science
ETH Zurich
8092 Zurich, Switzerland

[2]Electron Microscopy
ETH Zurich
8093 Zurich, Switzerland

Abstract

Automatic, defect tolerant registration of transmission electron microscopy (TEM) images poses an important and challenging problem for biomedical image analysis, e.g. in computational neuroanatomy. In this paper we demonstrate a fully automatic stitching and distortion correction method for TEM images and propose a probabilistic approach for image registration. The technique identifies image defects due to sample preparation and image acquisition by outlier detection. A polynomial kernel expansion is used to estimate a non-linear image transformation based on intensities and spatial features. Corresponding points in the images are not determined beforehand, but they are estimated via an EM-algorithm during the registration process which is preferable in the case of (noisy) TEM images. Our registration model is successfully applied to two large image stacks of serial section TEM images acquired from brain tissue samples in a computational neuroanatomy project and shows significant improvement over existing image registration methods on these large datasets.

1. Introduction

The registration of images models the geometrical back transformation that maps each image to the next and to the previous image in a sequence, in order to resolve correspondences between images. In medical image analysis for example, image registration is used to identify anatomical structures of a patient according to a medical atlas, or to compare images from one patient taken at different times. In computer vision registration is required as a preprocessing step for motion tracking and 3D reconstruction. In this paper we address the registration of electron microscopy images for geometry extraction, which poses a challenging new task due to a low signal to noise ratio and structural changes between images.

In order to record a specimen with a transmission electron microscope (TEM), the specimen (conserved in a paraffin block) has to be cut into ultra thin sections of about 70 nm thickness. These sections are recorded with the microscope, stitched together, registered to form a 3D image stack and eventually segmented to gain a 3D reconstruction of the specimen. Today most of this work involves time consuming manual labor, which renders the process very tedious for the microscopist as well as quite prone to errors [8].

In this paper we present an automatic approach for the registration related part of the 3D reconstruction pipeline. The most challenging part consists of the non-linear registration of images from single sections. Transmission electron microscopy with its special preparation of biological samples causes problems that are not solved by existing image registration methods. The problems are related to three major error sources,

(i) The electron beam of the microscope causes tissue damage in the specimen and, therefore, they can only be exposed to a limited electron dose. The image quality is rather poor, i.e. they have low contrast and a low signal to noise ratio.

(ii) In addition the preparation process often causes artifacts in the images that prevented an automation of the image processing pipeline so far. Examples of such artifacts are cuts in the samples caused by the cutting of the 70 nm thin slices or staining (biological markers) errors which result in darker blobs in the image.

(iii) The image content changes strongly between two slices. Correspondences are still recognizable by human experts, but existing registration methods fail to detect these semantic relations.

In addition to the above mentioned deformations caused by the biological sample preparation, the electron microscope introduces a significant amount of non-linearity by non-linear lens distortions. Therefore in this paper we provide two methods to correct for non-linear deformations:

978-1-4244-2243-2/08/$25.00 ©2008 IEEE

(i) an auto-calibration method to compensate for non-linear lens distortions that are the same in all images, (ii) a fully automatic approach to non-linear registration of TEM images between different sections. Both problems can be solved by regression of correspondence points. While the distortion correction works on images taken from the same section, the non-linear registration method has to deal with structural changes between physical sections and therefore has to be much more robust than the non-linear lens correction.

We solve the latter problem by way of an expectation maximization algorithm that calculates a non-linear warping which is parametrized by a polynomial kernel expansion of reference points. The correspondence points are not a priori fixed but selected during the registration process. Anomalies, which are caused by the biological sample preparation, are estimated in the image. The study clearly demonstrates that the estimation of image anomalies substantially improves the registration process, which can be measured by visual inspection as well as by a statistical evaluation on a large dataset of TEM images.

An overview of image registration is given in [19] in general and in [7] for medical image analysis. In [5, 9, 11, 15] image registration methods are introduced that incorporate both spatial distance of correspondence points as well as intensity values. Gay-Bellile et al. registered images by using thin-plate splines [4]. Thin-plate splines were also used for point matching [3] which is highly related to image registration [16]. A robust framework to estimate optical flow was proposed in [1, 2], which is also related to image registration. An iterative approach to register TEM images of neuronal structures based on Gabor features is presented in [10]. Luther et al. [13] documented damages of the specimen caused by the electron beam of the TEM that leads to non-linear transformations of the acquired image. In our work we extend the previous work by introducing additional "visibility" variables that detect image anomalies. This concept is related to identifying regions that are visible from both images in stereo reconstruction. E.g., Strecha et al. [17, 18] used hidden visibility variables to detect visible and non-visible regions.

2. Stitching, Calibration and Correction of Nonlinear Distortions

In electron microscopy it is common practice to record a large region by taking several translated overlapping images. Until today the resulting stitching problem is solved manually [8]. We propose an automatic landmark extraction approach based on SIFT features [12] to extract correspondence points from overlapping areas and solve the stitching problem automatically. SIFT features are calculated for each image to be stitched. Then a nearest neighbor search finds corresponding pairs of feature vectors, between images of the same section, that do not exceed a given distance threshold.

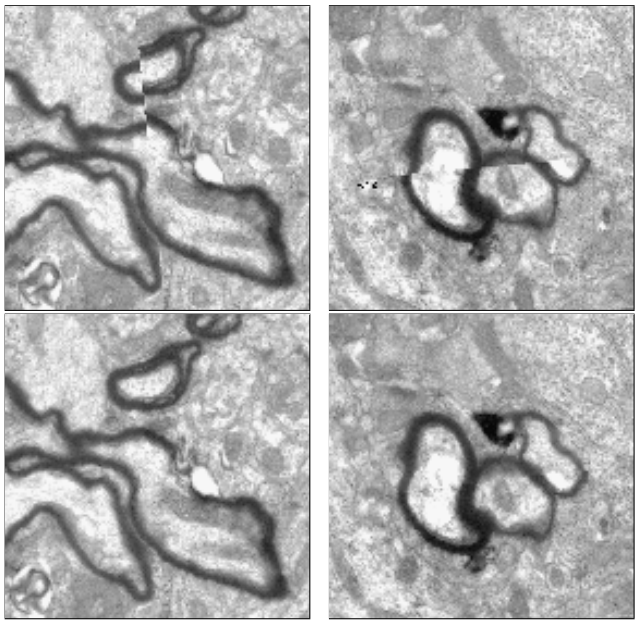

Figure 1. Two example regions of the stitching intersection. In the top row without distortion correction the image border is clearly visible. In the second row the distortion correction makes a seamless stitching possible.

As can be seen in Figure 1, an affine transformation is not sufficient to stitch images that contain lens distortions caused by the electron microscope. A non-linear correction is required to stitch the images neatly together. Here we introduce a new method to estimate the non-linear transformation that corrects the image deformations. Note that the same non-linear transformation is applied to every single image for correction. Thus, only the image acquisition process can be the cause for this distortion.

Our approach estimates a rotation matrix $R^{(i)}$ and a translation vector $T^{(i)}$ for each single image i and a non-linear distortion α that globally corrects all images. For the non-linear transformation we use an explicit polynomial kernel expansion to map the points x_i into a higher dimension

$$\phi(x_i) = [1, x_{i1}, x_{i2}, x_{i1}^2, x_{i1}x_{i2}, x_{i2}^2, \ldots, x_{i2}^d]^T \quad (1)$$

The transformation matrix α then projects these points back into the image plane, leading to a nonlinear transformation.

The non-linear transformation should correct for errors of all correspondence points that cannot be corrected by the individual rigid transformations alone. This extension leads to the following optimization term:

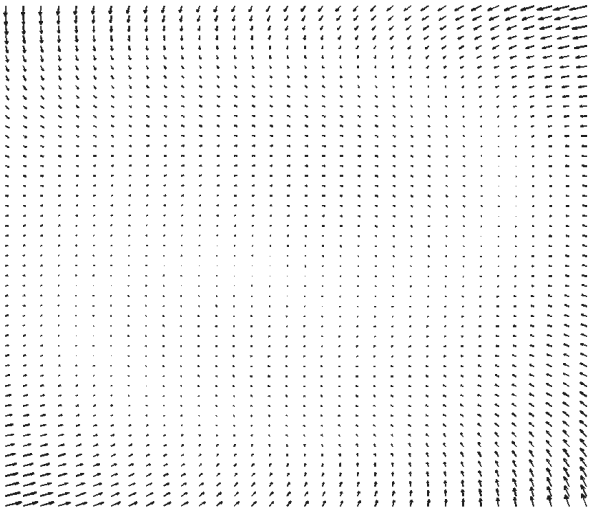

Figure 2. Needle diagram of the non-linear distortion correction that was applied to all images

$$\min_{\alpha,R,T} \sum_{i}^{B} \sum_{j}^{B} \Big\| \Big(\phi(X^{(i,j)}) \alpha R^{(i)} + T^{(i)} \Big) - \Big(\phi(X^{(j,i)}) \alpha R^{(j)} + T^{(j)} \Big) \Big\|^2 + \lambda \sum_{i}^{B} \Big\| \phi(X^{(i)}) \alpha - X^{(i)} \Big\|^2 \quad (2)$$

$X^{(i,j)}$ denotes the correspondence points between image i and image j and $X^{(i)}$ denotes the points of image i. The second part of the sum is a regularization term, expressing the preference for the identity transformation as the natural choice for α. To avoid a strong bias, the regularization weight λ should be set to a rather small value.

In practice the term can be optimized by iteratively estimating first the rigid transformations $T^{(i)}$ and $R^{(i)}$ and then the distortion matrix α. So far however, one or two iterations were sufficient to derive α for all our image sets. Without any distortion correction the median error of the affine stitching amounts to 15.17 pixels. Our experiments show that with a distortion correction based on a 5 fold overlap the error drops below 2 pixels.

Figure 1 shows a stitching of distorted images before and after the calibration. Figure 2 depicts a needle diagram of the non-linear distortion applied to all images. As can be seen, the distortion is fairly small in the center of the image and becomes larger at the image borders.

3. Expectation-Maximization

After the preprocessing steps, the images now have to be aligned into one image stack. Due to cutting errors (cracks or folds), different tissue stress during the preparation and variations in slice thickness a non-linear transformation model is necessary to register the images in detail. Furthermore, due to the difficult preparation process, a lot of images show artefacts with high contrast caused by staining errors, that lack correspondences in the target image. Our aim is to exclude points located on these anomalous structures from the matching process.

The problem is modeled as a mixture model where we not only mark these non-relevant structures, but we also solve the whole correspondence problem during the registration process itself. The optimization is achieved by the expectation-maximization algorithm. In a Bayesian framework the optimal transformation matrix β maximizes the posterior probability

$$p(\beta|X,Y) = \sum_{M \in \mathcal{M}} \frac{p(X,Y|\beta,M) \cdot p(\beta) \cdot p(M)}{p(X,Y)}, \quad (3)$$

where X and Y correspond to the warp image and the target image. The variable M denotes a binary correspondence matrix. M_{ij} is one, if point x_i in the warp image corresponds to point y_j in the target image and zero otherwise. In addition a point x_i can be assigned to an outlier class, denoted as M_{i0}. Thus, the whole matrix M is of size $n_1 \times (n_2 + 1)$ where n_1 is the number of chosen points in the warp image and n_2 is the number of possible correspondences for each of these points in the target image. The set $\mathcal{M} = \{0,1\}^{n_1 \cdot (n_2+1)}$ denotes all admissible assignment matrices M.

Application specific constraints on the assignments can be modeled by an appropriate definition of $\mathcal{M}$. As each point x_i should be assigned to only one correspondence point y_j we define $p(M)$ to be zero for configurations that assign more than one correspondence point to x_i and assume all valid configurations to be uniformly distributed:

$$p(M) \sim \begin{cases} 1 & \text{if } \sum_{j=0}^{n_2} M_{ij} = 1 \text{ for all } i = 1, \ldots, n_1 \\ 0 & \text{else} \end{cases} \quad (4)$$

This definition of $p(M)$ ensures that each point x_i is assigned exactly to one correspondence point or is marked as not relevant.

To define $p(\beta)$ the components of the solution vector β are assumed to be normally distributed. Therefore, we introduce a ridge penalty which is described by the normal distribution $\varphi_{\mu,\sigma}$ with location parameter μ and variance σ^2 as prior distribution

$$p(\beta) = \prod_{i=1}^{n_\beta} \varphi_{0,1/\sqrt{\lambda}}(\beta_i) \quad (5)$$

for β. The parameter λ is the ridge penalty that controls the complexity of the regression function.

The distribution of $p(X, Y|\beta, M)$ should depend on the similarity of the correspondence points based on gray values, as well as on the quality of the geometric fit. Furthermore, we need to take care of image anomalies that are assigned to the outlier class. This outlier class is modeled as a uniform distribution. The complete data likelihood is distributed as

$$\begin{aligned} p(X, Y|\beta, M) \quad \sim \quad & \prod_{i=1}^{n_1}\prod_{j=1}^{n_2}\Big(\varphi_{0,\sigma_1}\big(v(x_i) - v(y_j)\big) \cdot \\ & \varphi_{0,\sigma_2}\big(\phi(x_i)\beta - y_j\big)\Big)^{M_{ij}} \\ & \cdot \prod_{i=1}^{n_1}\big(\varphi_{0,\sigma_1}(c\sigma_1)\big)^{M_{i0}} \end{aligned} \tag{6}$$

Here $v(x_i)$ is a vector of the gray values of a small patch centered at x_i. The difference of two such patches serves as a dissimilarity measure that is easy to compute and takes context information about a small area around the points into account. The non linear transformation is modeled again as a polynomial kernel expansion (Eq. 1) followed by a multiplication with the matrix β to project the points back to two dimensions.

The second factor provides a penalty for points that are marked as not relevant. The constant c is given as the 0.98 quantile of the cumulative chi square distribution, where the degrees of freedom in principle correspond to the number of pixels in $v(x_i) - v(y_j)$. To reduce the influence of noise in the similarity measure, we perform a principal component analysis and project the high dimensional difference vectors down to the eigenvectors that correspond to the largest 98% of the eigenvalues. This also reduces the degree of freedom of the cumulative chi square distribution. Thus, the outlier factor transforms the assumed normal distribution of the gray value similarity into a heavy tailed distribution, providing a robust solution for outliers caused by non relevant elements.

In order to maximize $p(\beta|X, Y)$ which yields an optimized transformation, we maximize the logarithm of $p(X, Y|\beta, M) \cdot p(\beta)$ under the constraint that M is a valid matrix, i.e. $\sum_{j=0}^{n_2} M_{ij} = 1$ for all $i = 1, \ldots, n_1$. Since the assignment variables M are unobservable, we use the EM-algorithm to maximize the joint log-posterior. The algorithm iterates between estimating the expectation of the latent variables M_{ij} while keeping β fix and maximizing the joint log-posterior while keeping the expectation values of M constant. The variances for the normal distributions are also calculated during the maximization step. The log-posterior is maximized with respect to the transformation β as well as the variances of the normal distributions σ_1 and σ_2.

E-step: In each iteration the variables M_{ij} are replaced by their conditional expectation given β. The expectation values are calculated using the currently optimized β. Under the condition that M is a valid assignment matrix, we derive the following result:

$$\begin{aligned} \gamma_{ij} \quad &= \quad \mathbf{E}[M_{ij}|X, Y, \beta] \\ &= \quad \frac{p(X, Y|\beta, M_{ij} = 1)}{\sum_{l=0}^{n_2} p(X, Y|\beta, M_{il} = 1)} \end{aligned} \tag{7}$$

M-step: The expectation of the joint log posterior has the same form as the joint log posterior itself, but with M_{ij} replaced by γ_{ij}. The parameter $\beta, \sigma_1, \sigma_2$ are then computed by maximizing the expectation of the joint log posterior. For the transformation β this MAP approach yields a weighted ridge regression problem [6] with weights γ_{ij}. The transformation matrix β is maximized by

$$\beta \leftarrow (\phi(\widetilde{\mathbf{X}})^T \mathbf{\Gamma} \phi(\widetilde{\mathbf{X}}) + 2\lambda \mathbf{I})^{-1} \phi(\widetilde{\mathbf{X}})^T \mathbf{\Gamma} \mathbf{Y} \tag{8}$$

where Γ is a $(n_1 \cdot n_2) \times (n_1 \cdot n_2)$-dimensional diagonal matrix of the weights γ_{ij}. The $(n_1 \cdot n_2) \times 2$ matrix $\widetilde{\mathbf{X}}$ contains n_2 copies of each position vector x_i and the $(n_1 \cdot n_2) \times 2$ matrix $\mathbf{Y}$ contains n_2 possible correspondence points for each position x_i. The parameter λ is the regularization parameter defined by the prior distribution $p(\beta)$ (Eq. 5). In our experiments $\lambda = 0.001$ sufficiently regularizes the assignments.

The standard deviations are updated by

$$\sigma_1 \leftarrow \sqrt{\frac{\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\gamma_{ij} \cdot \rho(x_i, y_j)^2}{\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\gamma_{ij}}} \tag{9}$$

$$\sigma_2 \leftarrow \sqrt{\frac{\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\gamma_{ij} \cdot ||\phi(x_i)\beta - y_j||^2}{\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\gamma_{ij}}} \tag{10}$$

where σ_1 and σ_2 are invariant to outliers since $\gamma_{i,0} \approx 1$ for these points and therefore $\gamma_{i,j} \approx 0$ for $1 \leq j \leq n_2$.

Choice of initial points x_i: So far all warping points x_i are assumed to be arranged on a regular grid. While this design ensures that all interesting structures in the image are covered by a warp point nearby, interest points are often placed in background areas. To increase precision we would like to position each point directly in content rich parts of the image while still covering all biologically relevant structure in the image. For this purpose we calculate the entropy of the intensity value in a 13^2 neighborhood around each point in the image. The entropy is high for pixels along structures with a high contrast. Now we shift each warp point x_i of the regular grid to the position with the highest entropy value in its neighborhood. This local adaptation

method preserves the coverage of the whole image while emphasizing areas with rich image content.

4. Data and Experiments

For our experiments we have used images gathered in a computational neuroscience project. When imaging with a TEM it must be possible for single electrons to penetrate the probe. Therefore the specimen is first stained, then embedded into resin and cut into ultra thin sections of 70 nm thickness.

During this whole process the three major steps that may cause artefacts in the image are the staining, the cutting and the recording with the electron beam. Staining may produce additional dark areas in the image that do not correspond to original biological structures. During the cutting process the slice is exposed to significant stress and it may be non-linearly transformed or it even can encounter fractures. Finally, exposure to an electron beam causes a mass loss of the specimen and leads to additional transformations.

Figure 3 depicts an example image. On the left side is an image of typical quality. The structure in the left bottom quarter of the image with the pike on the right hand side shows a dendrite. The smaller ellipse like structures over the image are myelinated axons. The four smaller images on the right show examples of image artefacts, that are caused by the sample preparation: In the lower right corner we see an example of a crack in the specimen caused by the cutting process. The dark spots in the upper left image and the dark stripes in the upper right and lower left image are artefacts of the staining procedure.

We tested our approach on two series of electron microscopy images. The first series contains 97 images that were taken at 3400x magnification with a resolution of 1032x1376 pixels, one image per section. The second series consists of 284 images taken at 13500x magnification that were distortion corrected and stitched into 71 section images. The resolution of the second stack is 2672x4008 pixels.

In a preprocessing step we correct the radial illumination gradient visible in the images by dividing each image with a smoothed version of itself ($\sigma = 30$ pixel). Then an initial affine transformation based on SIFT features is estimated for each image. This transformation is then refined with our warping approach based on expectation maximization. For comparison we also provide the refinements for ridge regression with the ordinary least squares error

$$\min_{\beta} E(\beta) = \min_{\beta} \sum_{i=1}^{n} \|\phi(x_i)^T \beta - y_i\|^2 + \lambda \|\beta\|^2 \quad (11)$$

and the robust version, where the Huber loss function is used instead of the least squares error

$$L_c(\xi) = \begin{cases} c|\xi| - \frac{c^2}{2} & \text{for } |\xi| > c, \\ \frac{\xi^2}{2} & \text{for } |\xi| \leq c. \end{cases}$$

This robust method can be seen as a non-linear extension of the rigid approach described in [14]. As these methods need fixed corresponding points, we calculate the correlation coefficient of a patch around x_i and the appropriate patches of the target image. The result is then weighted with a Gaussian density centered at the position of x_i and a standard deviation of two times the standard error of the affine match. Each mapping point x_i is then assigned to its correspondence point y_i according to the maximum of the obtained function.

To provide a visual impression of the obtained warpings we first show difference images for the affine initialization and two warpings (Figure 4). The original images are shown in Figure 5. The darker the color in the difference image, the larger the absolute difference in gray values between the warped image and the target image. The images in the left column show the difference map for the robust affine transformation, the images in the middle the result for ridge regression with polynomial basis functions. The third column shows the difference maps for our new expectation-maximization method. If one focuses on the images in the upper row on the top left region as well as on the right and bottom border, one can observe dark stripes in the affine transformation image that are getting thinner for the least squares and even more thinner for the expectation maximization method. This error measure shows clearly that important edges are not matched very well by the affine transformation, but for the expectation maximization solution, there are only very small differences left over. Edges in the image are fitting significantly better than for the standard technique. In addition we tried to compare our method to optical flow methods [1, 2]. We could not find parameter settings that were able to cope with the significant structural changes between sections.

Since the detection and localization of image anomalies is important for our method, we show examples of the estimated image anomaly regions in Figure 5. The darker the color in this image, the more likely the region belongs to an artefact. The dark strip over the upper image as well as the dark blobs in the lower image are clearly detected as not relevant structures which is our desired goal. The information about these anomalies can now be used for further processing steps in computational neuroanatomy, e.g. in the 3D reconstruction of the neural connectivity structure. To visualize the estimated non-linear transformation of the image, we show a needle diagram of the transformation in Figure 6.

To demonstrate the improvement of the new method, we have registered both stacks of TEM images from the above

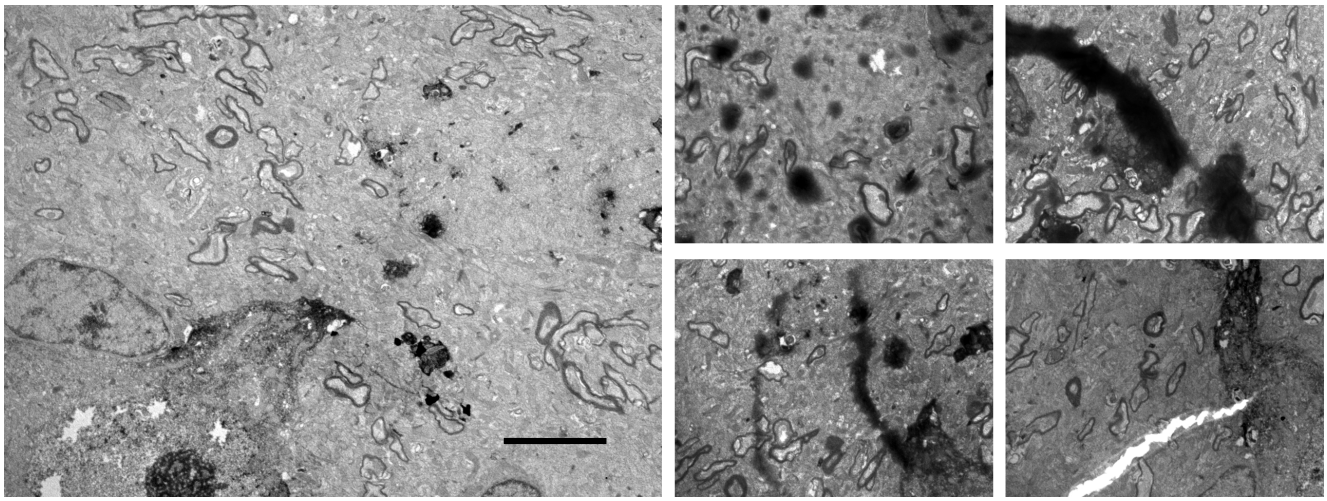

Figure 3. Examples of experimental data. On the left an image of typical contrast without major artefacts. The scale bar corresponds to $4\mu m$. On the right side some examples of artefacts caused by the preparation process.

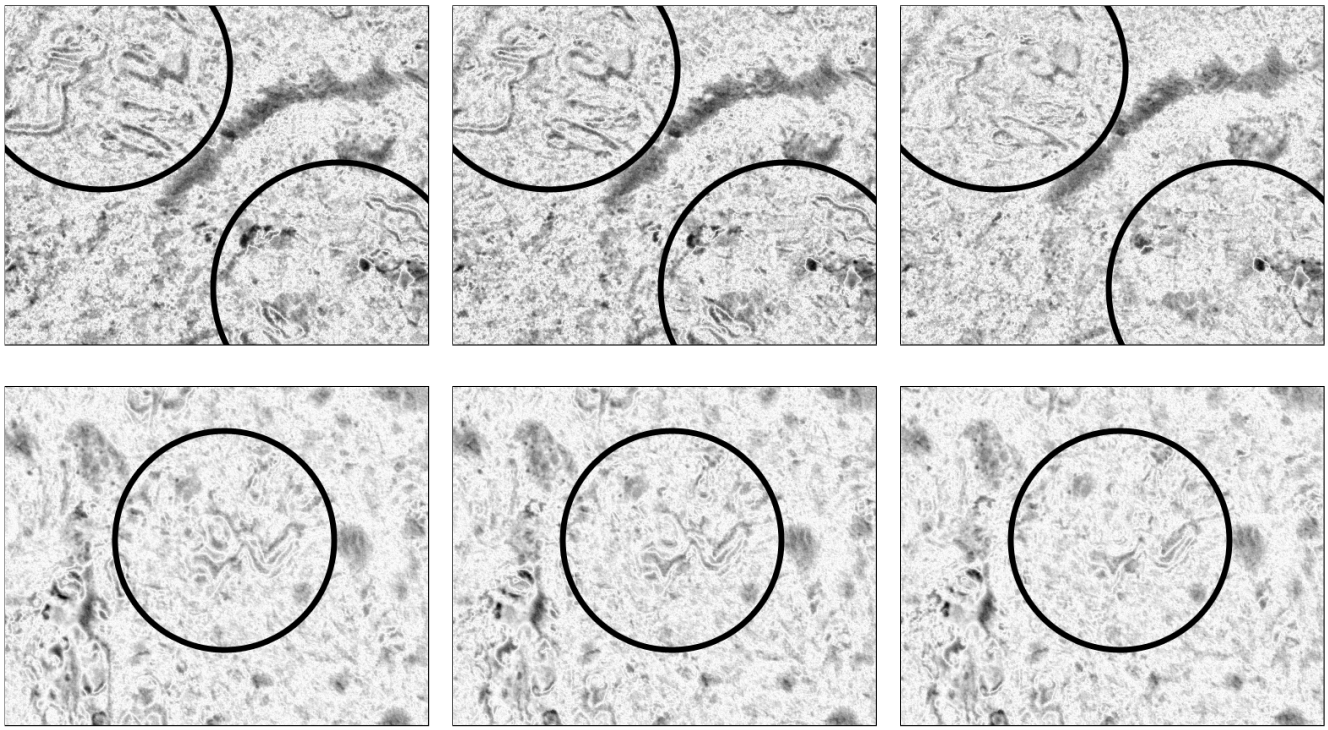

Figure 4. Difference images for (left) affine transformation, (middle) least squares matching with polynomial basis functions, (right) expectation-maximization including visibility estimation. The original images are shown in Figure 5

described neuroanatomy project and measured the cross-correlation between the target image and the warped image. Figure 7 shows the results for the two image stacks. Our EM approach has been initialized with the robust ridge regression solution. Therefore, improvements in cross-correlation values over the robust ridge regression solution serves as a measure of success for our model of image registration, i.e., large differences in cross-correlations denote a significantly better registration of the TEM images than with the robust Huber loss. The EM method clearly outperforms the other techniques by up to 15 percent gain in cross-correlation.

5. Conclusion

Registration of images is an important step in the 3D reconstruction. Especially in Transmission Electron Microscopy of biological samples, image anomalies occur fre-

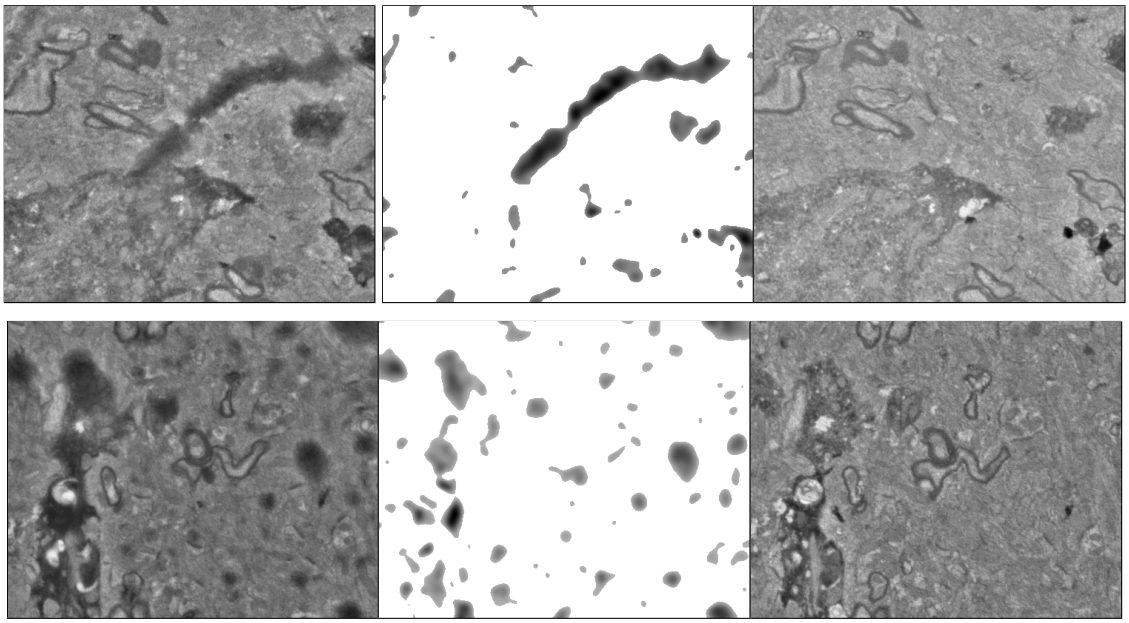

Figure 5. Left: The first image, Middle: The visibility map (Detected image anomalies). Right: The second image.

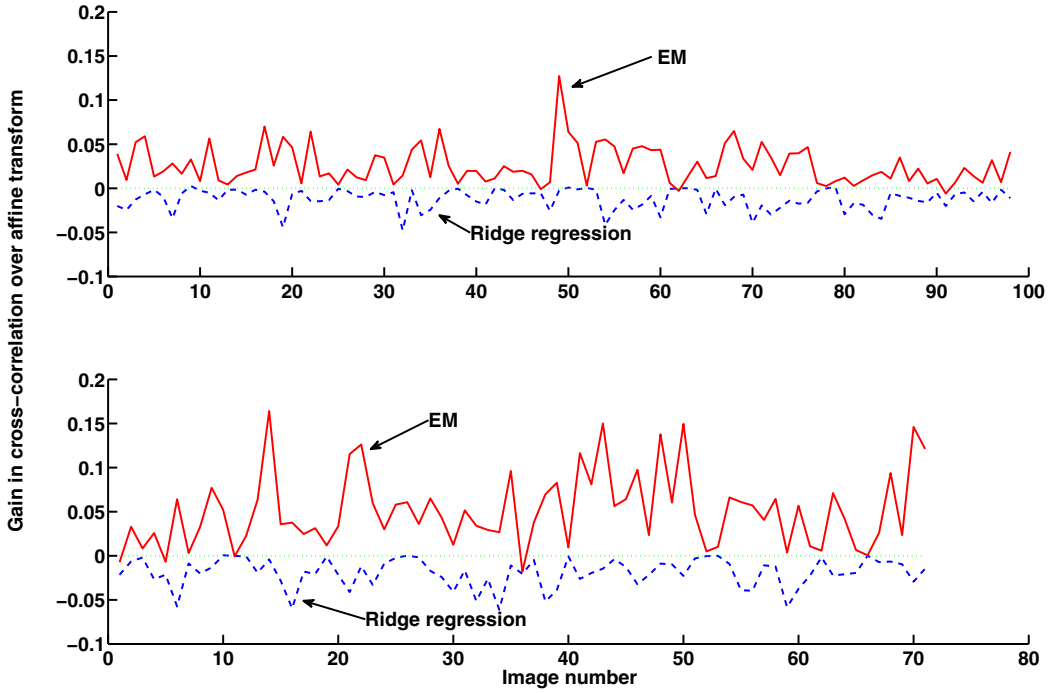

Figure 7. Gain or loss of EM warping (solid line) and ridge regression (dashed line) relative to robust Huber loss estimation. Top: the single image stack containing 97 images, bottom: distortion corrected and stitched image stack containing 71 images.

quently caused by the sample preparation process and by the image acquisition process. In this paper we propose a novel method for image registration that jointly estimates image anomalies and an image matching in a Bayesian model. The mixture model enables us to estimate assignment probabilities as well as probabilities for damages. The method performs superior to standard methods like linear affine transformation, and non-linear transformation. Even

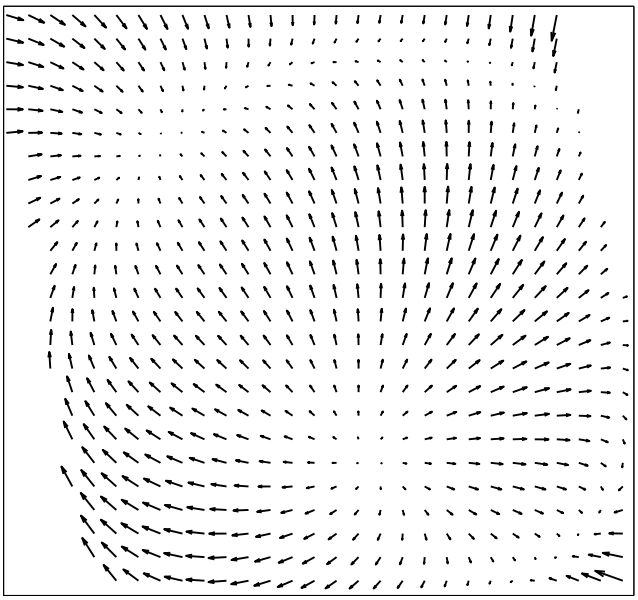

Figure 6. A needle diagram of non-linear image warping.

state-of-the-art outlier detection methods are inferior in performance compared to our mixture model. The expectation-maximization algorithm optimizes the model efficiently and is straight forward to implement. The experiments convincingly demonstrate that the model not only improves the image registration process, but that it also detects image anomalies. This quality of the algorithm will prove to be important for further steps in the automation of the microscopists tasks in image processing, like the 3D reconstruction of neuronal tissue.

Acknowledgement

We like to thank Nuno Maçarico da Costa and Kevan Martin, Institute of Neuroinformatics UNI-ETH Zurich, for valuable discussions on neuroanatomy and for providing the TEM images.

References

[1] M. J. Black and P. Anandan. The robust estimation of multiple motions: parametric and piecewise-smooth flow fields. *Comput. Vis. Image Underst.*, 63(1):75–104, 1996. 2, 5

[2] T. Brox, A. Bruhn, N. Papenberg, and J. Weickert. High accuracy optical flow estimation based on a theory for warping. In T. Pajdla and J. Matas, editors, *European Conference on Computer Vision (ECCV)*, volume 3024 of *LNCS*, pages 25–36, Prague, Czech Republic, May 2004. Springer. 2, 5

[3] H. Chui and A. Rangarajan. A new algorithm for non-rigid point matching. In *CVPR*, volume 2, pages 44–51, 2000. 2

[4] V. Gay-Bellile, M. Perriollat, A. Bartoli, and P. Sayd. Image registration by combining thin-plate splines with a 3D morphable model. In *ICIP*, pages 1069–1072, 2006. 2

[5] B. Glocker, N. Komodakis, N. Paragios, G. Tziritas, and N. Navab. Inter and intra-modal deformable registration: Continuous deformations meet efficient optimal linear programming. In *Information Processing in Medical Imaging*, volume 4584 of *Lecture Notes in Computer Science*, pages 408–420. Springer, 2007. 2

[6] T. Hastie and R. Tibshirani. *The Elements of Statistical Learning*. Springer, 2001. 4

[7] D. L. G. Hill and P. G. Batchelor. Medical image registration. *Physics in Medicine and Biology*, 46(3), 2001. 2

[8] B. K. Hoffpauir, B. A. Pope, and G. A. Spirou. Serial sectioning and electron microscopy of large tissue volumes for 3D analysis and reconstruction: a case study of the calyx of held. *Nature Protocols*, 2(1):9–22, 2007. 1, 2

[9] H. J. Johnson and G. E. Christensen. Consistent landmark and intensity based image registration. *IEEE Trans Med Imaging*, 21(5):450–461, 2002. 2

[10] P. König, C. Kayser, V. Bonin, and R. P. Würtz. Efficient evaluation of serial sections by iterative Gabor matching. *Journal of Neuroscience Methods*, 2001. 2

[11] W. Li and H. Leung. A maximum likelihood approach for image registration using control point and intensity. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(8):1115–1127, 2004. 2

[12] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 2004. 2

[13] P. K. Luther, M. C. Lawrence, and R. A. Crowther. A method for monitoring the collapse of plastic sections as a function of electron dose. *Ultramicroscopy*, 24(1):7–18, 1988. 2

[14] S. Ourselin, A. Roche, G. Subsol, X. Pennec, and N. Ayache. Reconstructing a 3D structure from serial histological sections. *Image and Vision Computing*, 19(1-2):25–31, January 2001. 5

[15] D. Perperidis, R. H. Mohiaddin, and D. Rueckert. Spatio-temporal free-form registration of cardiac MR image sequences. *Medical Image Analysis*, 9(5):441–456, October 2005. 2

[16] C. V. Stewart, C. Tsai, and B. Roysam. The dual-bootstrap iterative closest point algorithm with application to retinal image registration. *IEEE Transactions on Medical Imaging*, 22:1379–1394, 2003. 2

[17] C. Strecha, R. Fransens, and L. V. Gool. Wide-baseline stereo from multiple views: a probabilistic account. In *CVPR*, volume 2, pages 552–559, 2004. 2

[18] C. Strecha, T.Tuytelaars, and L. V. Gool. Dense matching of multiple wide-baseline views. In *ICCV*, volume 2, pages 1194–1201, 2003. 2

[19] B. Zitova and J. Flusser. Image registration methods: a survey. *Image and Vision Computing*, 21(11):977–1000, October 2003. 2