



## Fully automatic stitching and distortion correction of transmission electron microscope images

Verena Kaynig<sup>a,c,\*</sup>, Bernd Fischer<sup>b</sup>, Elisabeth Müller<sup>c</sup>, Joachim M. Buhmann<sup>a</sup>

<sup>a</sup> Department of Computer Science (D-INFK), ETH Zurich, 8092 Zurich, Switzerland

<sup>b</sup> European Molecular Biology Laboratory Heidelberg (EMBL), 69117 Heidelberg, Germany

<sup>c</sup> Electron Microscopy ETH Zurich (EMEZ), 8093 Zurich, Switzerland

### ARTICLE INFO

#### Article history:

Received 8 January 2010

Received in revised form 23 April 2010

Accepted 25 April 2010

Available online 5 May 2010

#### Keywords:

Automatic stitching

Auto-calibration

SIFT features

Robust transformation

Transmission electron microscopy

Image distortion

Electromagnetic lenses

### ABSTRACT

In electron microscopy, a large field of view is commonly captured by taking several images of a sample region and then by stitching these images together. Non-linear lens distortions induced by the electromagnetic lenses of the microscope render a seamless stitching with linear transformations impossible. This problem is aggravated by large CCD cameras, as they are commonly in use nowadays. We propose a new calibration method based on ridge regression that compensates non-linear lens distortions, while ensuring that the geometry of the image is preserved. Our method estimates the distortion correction from overlapping image areas using automatically extracted correspondence points. Therefore, the estimation of the correction transform does not require any special calibration samples.

We evaluate our method on simulated ground truth data as well as on real electron microscopy data. Our experiments demonstrate that the lens calibration robustly corrects large distortions with an average stitching error exceeding 10 pixels to sub-pixel accuracy within two iteration steps.

© 2010 Elsevier Inc. All rights reserved.

## 1. Introduction

The field of electron microscopy has recently achieved substantial advances in imaging of large samples. In neuroanatomy, mega pixel images are regularly captured to reconstruct the three dimensional connectivity between neurons (Dauguet et al., 2007; Akselrod-Ballin et al., 2009). Frequently, microscopists are interested in analyzing regions of a sample that are too large to be captured by a single image at a sufficient magnification. Therefore, it is common experimental practice to capture several translated images of such a large region of interest and to assemble these images afterwards to cover the whole area under investigation. This process is commonly referred to as stitching, mosaicing or montaging. When stitching two images together, two conditions have to be fulfilled:

1. a number of reliably recognizable contrast patterns have to be identified and captured by both images;

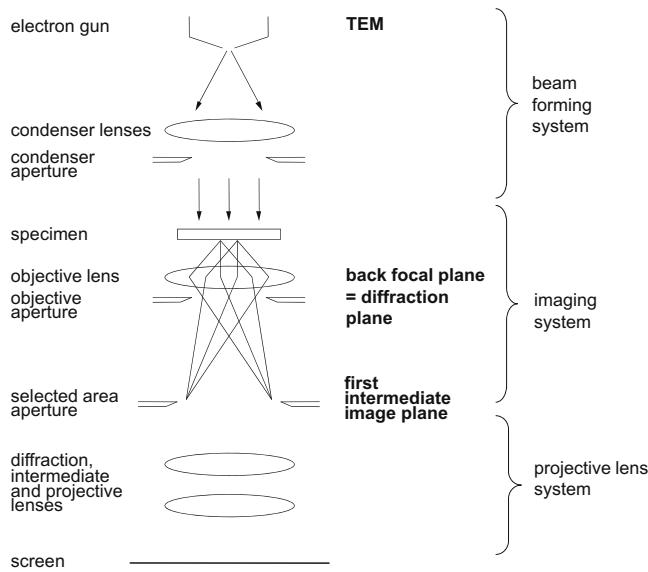
2. the images have to be free of distortions, e.g. perspective distortions or vignetting.

When substantial image distortions occur in the imaging process, it is difficult to seamlessly stitch the images together in a reliable way. Furthermore, distortion fields are growing with increasing distance to the image center. This observation suggests, that the distortion problem is less severe for CCD cameras with a small number of pixels than for image capturing devices with a large field of view despite the increasing demand for stitching. Using films, film plates or the latest generation of 10 mega pixel large-size CCD cameras, the image distortions can severely perturb reliable stitching of large field of view images.

Hardware adjustments, e.g. by changes of the electron optical conditions, would remove distortions in the most faithful and efficient way as it would restore the signal before analog to digital conversion and, therefore, it would recover the signal with the smallest possible loss. Such a strategy succeeds e.g. for light microscopes, where objective lens systems are built to minimize distortions. In electron microscopy, however, there exist several obstacles that render hardware controlled distortion corrections difficult. First, the distortions can sensitively depend on the electron optical configuration selected for image acquisition. This dependency implies that the correction would need to be adaptable to different parameter settings. Second, the distortion

\* Corresponding author at: Department of Computer Science (D-INFK), ETH Zurich, 8092 Zurich, Switzerland.

E-mail addresses: [verena.kaynig@inf.ethz.ch](mailto:verena.kaynig@inf.ethz.ch) (V. Kaynig), [bernd.fischer@embl.de](mailto:bernd.fischer@embl.de) (B. Fischer), [elisabeth.mueller@emez.ethz.ch](mailto:elisabeth.mueller@emez.ethz.ch) (E. Müller), [joachim.buhmann@inf.ethz.ch](mailto:joachim.buhmann@inf.ethz.ch) (J.M. Buhmann).



**Fig. 1.** Sketch of the basic components of a transmission electron microscope. Components influencing the distortions are the condenser lens system, the objective lens and the projective lens system.

parameters are commonly specific for the respective microscope. Third, little information has been reported in the literature about typical characteristics of imaging distortions induced by electron microscopes. In Fig. 1, a sketch of a transmission electron microscope (TEM) is shown. The main electron optical components of a TEM to be identified as possible sources for distortions are the condenser lens system, the objective lens and the projective lens system. The condenser lenses define the beam incident on the specimen, i.e. an almost parallel beam (cross-over far after the specimen), a focused beam (cross-over on the specimen) or an over-focused beam (beam cross-over before the specimen). The objective lens is located directly below the sample. This lens has short focal length and gives rise to a first magnified image. The first intermediate image plane is further magnified by the projective lens system, which consists of all lenses below the objective lens, e.g. diffraction, intermediate and projective lenses. Finally, the magnified image is captured by the camera or a viewing screen. The objective lens is well known for giving rise to distortions, because it suffers like every rotationally symmetric electromagnetic lens from spherical aberration. This effect mainly leads to a blurring of the images being equivalent to a loss of resolution. The spherical aberration exhibits its main degrading impact in high-resolution imaging. The projective lens system so far has not been thoroughly analyzed in the literature with respect to distortions. The large variety of magnifications accessible by the projective lens system is achieved by a combination of two effects. The strength of the lens current can be controlled for each individual lens separately, resulting in a change of the focal length of the respective lens. The experimenter can also select different combinations of the set of lenses below the objective lens. As charged particles move on a helical trajectory in an electromagnetic field, a change of the focal length of a lens not only causes the desired change in magnification, but it also induces a comparably small image rotation, while a different combination of the lenses can generate a large jump in image rotation. The condenser lens, the objective lens and the projective lens systems may all suffer from instabilities caused by hysteresis effects when changing the focal length of electromagnetic lenses. Our experiments clearly and unambiguously demonstrate that the dominant component of the distortion field is caused by the projective lens system. Since adjustment of

the projective lens system is not available to the microscopist, an improvement of the images is only possible by image processing after the image acquisition.

In this paper we present a fully automatic calibration and stitching approach, that can correct for non-linear distortions in the images caused by the electromagnetic lenses, while preserving structural information of the images. The approach has been implemented as an open source plugin for the [ImageJ Rasband, 1997–2010 distribution Fiji Fiji, 2010](#), is freely available and also has been integrated into [TrakEM2 Cardona, 2006](#). For a reliable estimate of the distortion correction we recommend to use an initial set of nine images in a  $3 \times 3$  grid structure with at least 50% overlap. The large overlap is essential to estimate the distortion correction for the central area of the field of view of the microscope and thus is only required for the nine calibration images. Afterwards, the estimated correction transformation can be applied to all images taken under equivalent electron optical conditions. For images with low contrast, like cryo TEM images, it is necessary to increase the number of calibration images as well as the size of the overlap region to ensure that the extracted correspondences are equally spread over the whole field of view of the microscope.

## 2. Related work

Lens calibration is commonly achieved by use of special calibration images, like pictures of checkerboards or other objects with straight lines ([Devernay and Faugeras, 2001](#); [Gremban et al., 1988](#); [Zhang, 2000](#)) in order to reproducibly correct the images after their acquisition. However, this method is not applicable to electron microscopy. While there are crystalline structures that might be used as calibration samples (i.e. a non-scale checker board), it is possible that the whole process of removing the calibration sample from the microscope, inserting the actual sample of interest and refocusing on that sample again changes the distortion field. This uncontrollable variability is caused by the electromagnetic lenses, which do not correspond to fixed shaped glass lenses of camera objectives. Instead they change shape due to different configurations of the microscope parameters. One important property of these electromagnetic lenses is that once a different parameter setting changed the distortion field, one cannot obtain the old distortion field just by going back to the old parameter setting, e.g. due to hysteresis effects.

Other methods often incorporate a specific distortion model, like radial symmetric “barrel” or “pincushion” distortions ([Wang et al., 2009](#); [Shih-Schon and Bajcsy, 2001](#); [Hartley and Kang, 2007](#)). While in principle, these distortions may affect imaging in electron microscopes, we prefer to not restrict our distortion model to specific shapes as done in [Sawhney and Kumar \(1999\)](#); [Stein \(1997\)](#). This generality enables superior correction in case of non-ideal electromagnetic fields in the microscope column. To our knowledge the only alternative approach to calibrate electron microscopes is described in [\(Koshevoy et al., 2006\)](#). Koshevoy et al. parametrize lens distortions by Legendre polynomials where intensity variance is used as a similarity measure. Our new proposed method distinguishes itself in three major points from the method by Koshevoy et al.: First, SIFT features are used to measure similarity between image patches/corners instead of intensity. This choice renders the approach more robust and therefore better applicable to handle noisy electron microscopy images than the alternative distortion correction. Second, our image assembly relies on a grid like matching, where we maximize the similarity between all overlapping pairs jointly. Instead, Koshevoy et al. only estimate a cascade of transformations over a predefined hierarchical order of the images. Due to the fact that errors are propagated over the image cascade, large errors can occur at the end nodes of such a cascade as a consequence of error propagation. Third, we

only correct a non-linear distortion field caused by the electron microscope that is shared by all images. Koshevoy estimates an additional non-linear transformation independently for each image. This procedure renders the approach problematic, as the non-linear correction induces structural changes to the images. For example, consider the assembly of images from two consecutive sections from a 3D specimen. The structural analysis should detect the structural variation between sections, but a non-linear transformation would compensate for it and thus suppresses the relevant information. Our method instead corrects only for the distortion field that appears in all images. As structural changes differ between the images they are not affected by this correction. Marsh automatically extracts correspondences in tomograms by cross correlation of subvolumes stating that, manual interaction is necessary to exclude false positive matches even for a large support of  $80 \times 80 \times 70$  voxels Marsh, 2007. SIFT features instead yield robust matches and are typically computed at an interpolated position with sub-pixel accuracy, which significantly outperforms localization based on cross correlation. Lawrence et al. (2006) describe non-linear lens distortions as an issue for three dimensional tomogram reconstructions and account for these distortions by extending the linear reconstruction method to a curvilinear model of cubic order. The distortion model described in this paper enables the correction of non-linear transformations of higher than cubic order while simultaneously preserving structural information.

### 3. Calibration and stitching

The goal of stitching is to assemble a set of images that are capturing a larger part of the sample and that have sufficient overlap (Fig. 2). When the images are distorted by the electromagnetic lenses of the microscope, it is not possible to find an affine transformation that stitches the images together without major errors. Fig. 3 shows a zoom in two different regions with large stitching errors by an affine transformation. The stitching boundary can easily be detected in both images. A non-linear correction is required to stitch the images neatly together. Here we introduce a new method to estimate this non-linear transformation that can be used to correct the images. Under the assumption that the non-linear distortion is only caused by the electron microscope, the same non-linear transformation is applied to every single image for correction. Fig. 2 illustrates the stitching scenario with distortions. The same region of the specimen is captured by two different images. The microscope induces the same distortion to all images. But, the original structure of the specimen is recorded in different areas of the local image coordinate system and thus the structure is dis-

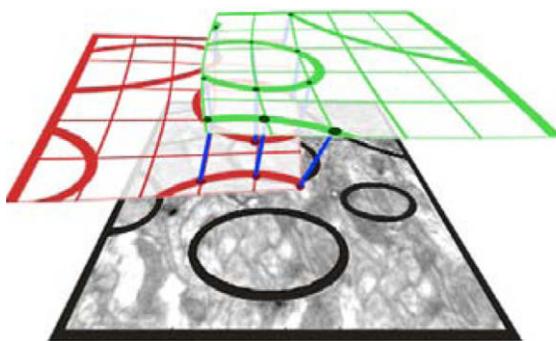
torted differently in both images. Our approach uses the information of the overlapping image areas to estimate a non-linear distortion correction that preserves the structural geometry of the images. For a series of images taken with the same parameter settings of the microscope, the distortion correction field has to be estimated only once in the beginning and can then afterwards be applied to all images. The method uses redundant information in overlapping images to estimate the non-linear transformation. Thus, no special calibration samples are required. The procedure can be divided into two steps:

1. extraction of correspondence points between image pairs,
2. joint calibration and stitching of images (iteration between auto-calibration and stitching).

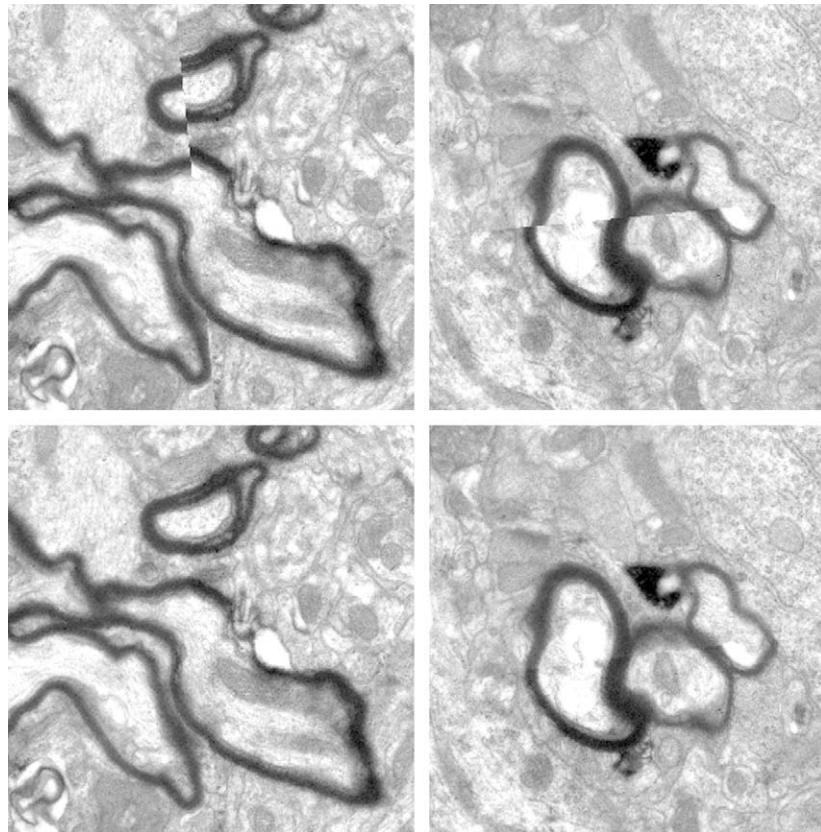
**Correspondence points:** Scale Invariant Feature Transform (SIFT) descriptors of local image properties were introduced by Lowe (2004) for fast scale invariant image matching. We use SIFT descriptors to extract correspondence points between overlapping image areas automatically. First, points of interest are detected in both images. These landmarks preferentially cluster around regions with high contrast. Then scale and rotation invariant feature descriptors based on local gradient information are calculated for each point. In the next step, correspondences are detected according to Euclidean distance in feature space: For each point  $x^{(i)}$  in image  $i$ , a correspondence point  $x^{(j)}$  in image  $j$  is identified that is the nearest neighbor of  $x^{(i)}$  in feature space. If this nearest neighbor point is significantly closer to  $x^{(i)}$  than all other points of image  $j$  in feature space, then the points  $x^{(i)}$  and  $x^{(j)}$  are marked as correspondent.

In the context of lens distortion correction in TEM images it is important, that the automatically extracted correspondences are robust with respect to non-linear transformations and noise in the images. In order to demonstrate the reliable performance of SIFT features in this context we applied a highly non-linear transform to a neuroanatomical TEM image. The results are shown in Fig. 4. The original image (a) and the transformed image (b) were normalized to a gray value range from zero to one and white noise with a standard deviation of 0.1 was added to both images. Subfigure (c) depicts correspondences found in the images by a line for each point pair. The line connects the coordinates of the point from the original image with the coordinates of the point from the transformed image when plotted into the same coordinate system. As can be seen, the whole image range is densely covered with correspondences. Furthermore, the needle diagram of the estimated correspondence connections forms a smooth field which follows the ground truth transformation (d).

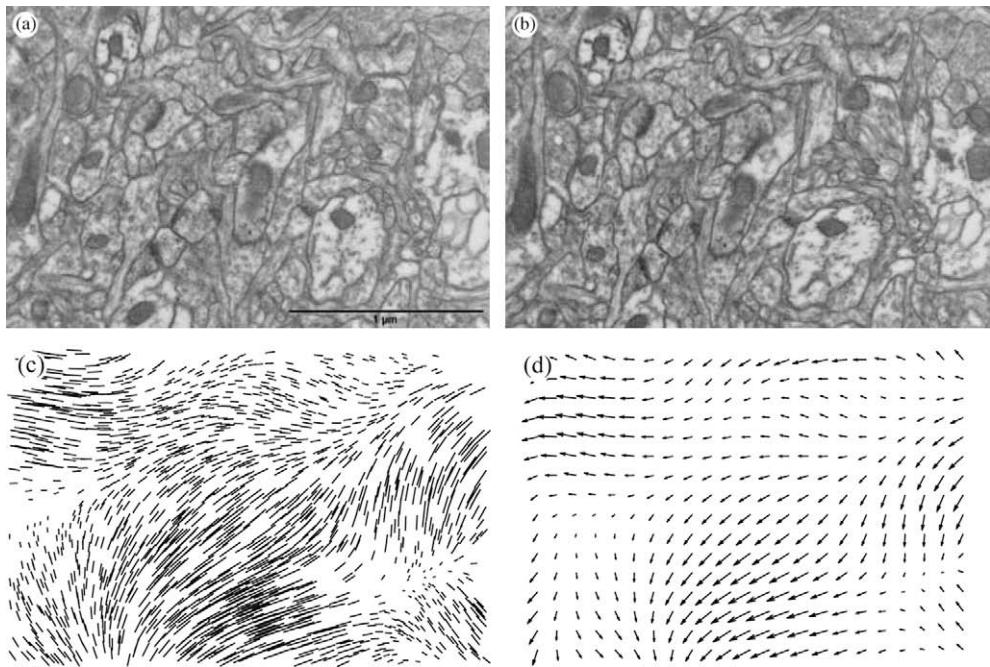
Cryo TEM images are challenging for correspondence point extraction due to low contrast, noise, and preparation artifacts. Fig. 5 depicts correspondences automatically extracted from two cryo TEM images. The left image shows a cryo section containing a plasma membrane, vesicles and a mitochondrion in the lower right corner. The image on the right shows the same section with a slightly translated field of view. Preparation artifacts in the form of compression waves and knife marks are clearly visible in both images. Despite the low contrast, the noise and the preparation artifacts, 14 correct correspondences are automatically found in the area with structural information. The two false positive matches can be filtered out automatically based on large affine alignment error. While the landmarks provide enough information to align the two images, the correspondences are too scarce to estimate a non-linear distortion correction without risking overfitting. The main problem is not the noise or the low contrast in the images, but the lack of texture information. Therefore, we recommend to use a sample region with abundant texture information to estimate the distortion correction. Afterwards the estimated transformation can be applied to images with little texture.



**Fig. 2.** Illustration of the stitching and calibration scenario. The same region of the specimen is covered by the overlapping image areas. Due to non-linear distortions in the images, overlapping regions cannot be matched with an affine transformation alone. The proposed self calibration method uses the redundant information in the overlapping areas to estimate the non-linear transformation field that corrects the distortion.



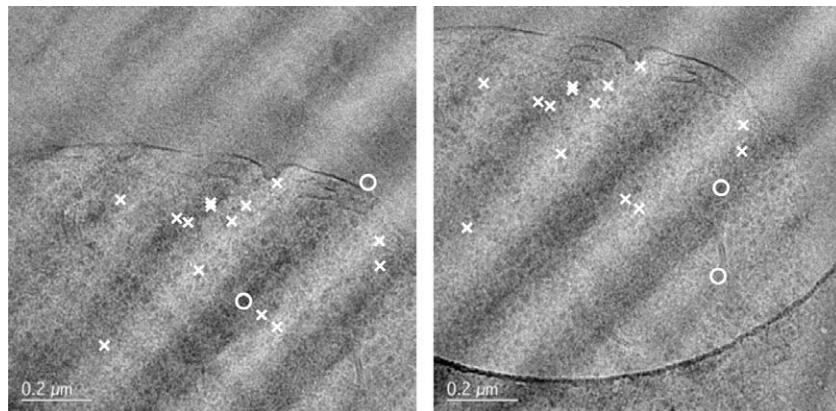
**Fig. 3.** Two example regions of the stitching intersection ( $2.7 \times 2.7 \mu\text{m}$ ). In the top line without distortion correction the image border is clearly visible. In the lower line the distortion correction produces a seamless stitching.



**Fig. 4.** SIFT correspondences for a highly non-linear deformation and considerable noise. (a) Original image with noise, (b) transformed image with noise, (c) connection between automatically obtained correspondences plotted in one coordinate system, (d) ground truth transformation.

In the following descriptions single points in image  $i$  are depicted with lower case letters  $x^{(i)}$  whereas upper case letters denote  $n \times m$  point matrices containing  $n$  points with  $m$  dimensions. Thus,  $X^{(i)} = \{x_1^{(i)}, \dots, x_n^{(i)}\}, x_l^{(i)} \in \mathbb{R}^2$  denotes the 2D-coordinates of  $n$  interest

points in image  $i$ . In order to remove potential outliers, we only take correspondences into account that are marked correspondent from image  $i$  to image  $j$  and vice versa. For all correspondence points between images  $i$  and  $j$ ,  $X^{(i,j)}$  contains the respective point coordinates



**Fig. 5.** Correspondences extracted from two translated cryo TEM images with low contrast and preparation artifacts in form of compression waves and knife marks. The sample contains a plasma membrane with an ice area above and structural information like vesicles and a mitochondrion below. The landmarks are concentrated on the area with textural information and neglect the ice part in the upper part of the images. Out of the 16 correspondences found, only two are outliers (marked with a circle).

in the coordinate system of image  $i$  and  $X^{(j,i)}$  contains the coordinates in the coordinate system of image  $j$ . Thus, the two sets  $X^{(i,j)}$  and  $X^{(j,i)}$  together build the set of correspondence points for images  $i$  and  $j$ .

*Stitching images:* stitching of images can be seen as an inversion of the coordinate transforms which are applied during the image acquisition process. The diagram in Fig. 6 illustrates the two main coordinate transforms which we consider in the stitching approach presented in this paper. The sample itself defines the original object coordinate system. If an image is taken from the sample, a region of interest from the original object coordinate system is mapped to its own image coordinate system. The transformation

between object coordinate system and image coordinate system is affine with a translation defined by the positioning of the field of view of the microscope, a rotation caused by the spiral movement of the electrons inside the microscope and a scaling according to the magnification settings. The lens distortion of the microscope now acts on the intermediate coordinate system and hence is the same for all images. Thus, for example the upper left corner of each image is distorted in the same way for all images, regardless of the absolute positioning of the region shown in the image in the original object coordinate system of the sample. To stitch a set of single images back to a common coordinate system, the two transformations applied to the images beforehand have to be inverted. Hence, first a non-linear transformation is applied to correct for the lens distortion and then the images are mapped with an affine transformation to the coordinate system of the stitched image. The coordinate system of the stitched image and the original object coordinate system of the sample are only equivalent up to an affine transform because the relative positioning of the sample in the object coordinate system is lost when the sample signal is mapped to the intermediate coordinate system (see Fig. 6). A common practice is to define one of the image coordinate systems as the stitching coordinate system and then map all other images to this coordinate frame.

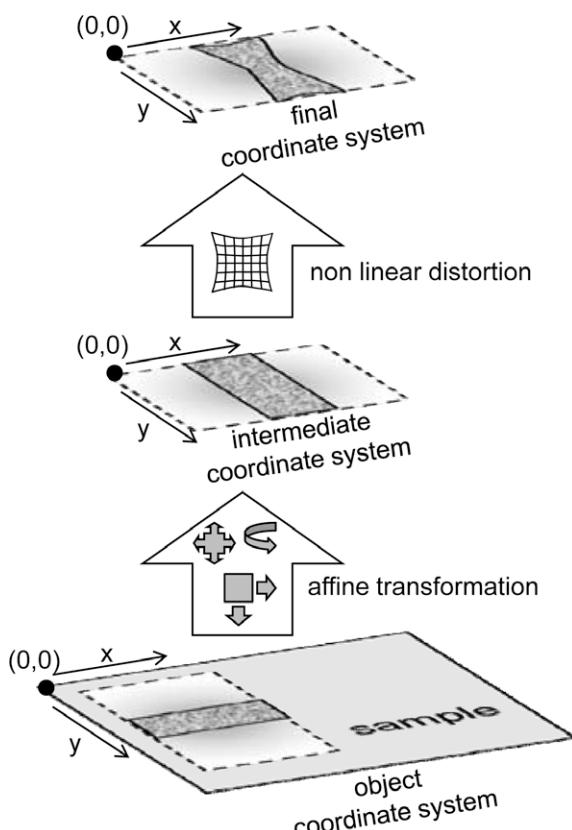
We now first consider the stitching problem with respect to the affine transformations and then extend the approach to the full task including the correction of lens distortions. Assuming that the microscope does not cause non-linear distortions, all images can be stitched together by estimating an affine transformation  $A^{(i)}$  for each image that transforms each image coordinate system to the common stitching coordinate system. The affine transformations are optimal, if the distances between correspondence points in the stitching coordinate system are minimal.

This translates to the following least squares problem:

$$\min_A \sum_{\substack{i,j=1 \\ j \neq i}}^B \|X^{(i,j)}A^{(i)} - X^{(j,i)}A^{(j)}\|^2 \quad (1)$$

The coordinates of the correspondence points  $X^{(i,j)}$  here are represented as homogeneous coordinates, transforming the 2D-coordinate  $(x,y)$  to a 3D-coordinate  $(x,y,1)$  by adding a one as the third dimension. The affine transformations  $A^{(i)}$  are then represented as  $3 \times 3$  transformation matrices.  $B$  denotes the total number of images to be stitched together.

As mentioned above, one affine transformation is set to identity to define the coordinate system of the stitching image. To estimate



**Fig. 6.** Illustration of the image acquisition process. First a region of interest from the object coordinate system is chosen and mapped with an affine transformation (translation, rotation, scaling) to its own intermediate coordinate system. The lens distortion then transforms this image further to the final image coordinate system.

the remaining affine transformations we define  $X^{(i)}$  as the set of all correspondence points in image  $i$ :  $X^{(i)} = \bigcup_{j=1}^B X^{(i,j)}$  and  $Y^{(i)}$  as the set of all corresponding points in the stitching coordinate system,  $Y^{(i)} = \bigcup_{j=1}^B X^{(j,i)} A^{(j)}$

The minimization problem (1) can now be solved by

$$A^{(i)} = \left( X^{(i)} \cdot X^{(i)} \right)^{-1} \cdot X^{(i)} \cdot Y^{(i)} \quad (2)$$

As  $Y^{(i)}$  depends on the affine transformations estimated, the problem is solved iteratively, alternating between updating the affine transformations  $A^{(i)}$  and the mapped correspondence coordinates  $Y^{(i)}$ .

The solution described so far weights the distance  $\xi$  between two correspondent points quadratically. But, SIFT features may lead to some false correspondence points, because points are only compared based on their feature vectors. Similarities in the structure or repeated patterns can cause incorrect correspondences. Thus, it is beneficial to use robust estimation methods, like Huber-loss or RANSAC to estimate the affine transformations (Kaynig et al., 2007; Saalfeld and Tomančák, 2008).

Applying Huber-loss, one replaces the squared loss function  $L(\xi) = \xi^2$  by the Huber-loss function which only gives linear weight to large errors:

$$L_c(\xi) = \begin{cases} c|\xi| - \frac{c^2}{2} & , \text{ for } |\xi| > c \\ \frac{\xi^2}{2} & , \text{ for } |\xi| \leq c \end{cases} \quad (3)$$

In Eq. (2) the Huber-loss introduces a diagonal weight matrix  $\Omega$  which reduces the influence of outliers on the estimated transformation accordingly.

$$A^{(i)} = \left( X^{(i)} \cdot \Omega \cdot X^{(i)} \right)^{-1} \cdot X^{(i)} \cdot \Omega \cdot Y^{(i)} \quad (4)$$

In the presence of non-linear lens distortions, affine transformations are not sufficient to reliably stitch the images together. A non-linear transformation is necessary to correct the images against distortions. To model the non-linear transformation we use an explicit polynomial kernel expansion to map the correspondence points into high dimensions and then estimate the transformation matrix  $\beta$  that projects the points back to the 2D image plane. The polynomial expansion of degree  $d$  for a 2D point  $x$  with coordinates  $(u, v)$  yields

$$\phi_d(u, v) = (1, u, v, u^2, uv, v^2, \dots, v^d). \quad (5)$$

For a  $n \times 2$  point matrix  $X$  the kernel expansion  $\phi_d(X)$  denotes the  $n \times \frac{(d+1)(d+2)}{2}$ -matrix where the kernel expansion is applied to all  $n$  rows separately. In all our experiments  $d = 5$  provided sufficient degree of freedom to estimate the non-linear distortion correction. As little is known about the characteristics of distortions induced by electromagnetic lenses, the polynomial kernel used for our experiments does not restrict the estimated transformation to correct for particular models like barrel or pincushion distortions. However, the method described in this paper can easily be restricted to specific transformation models by choosing another kernel function  $\phi$ , and lifting the points e.g. to a circular feature space (Geyer and Daniilidis, 2001; Claus and Fitzgibbon, 2005).

The non-linear transformation of a set of points  $X^{(i)}$  can now be written as a matrix multiplication  $\phi_d(X^{(i)})\beta$  with a  $\frac{(d+1)(d+2)}{2} \times 2$  transformation matrix  $\beta$  which projects the kernel expanded points back into 2D space.

The joint calibration and stitching can now be formulated as a minimization problem which directly follows the inverse transformations of Fig. 6: for each correspondence point the non-linear transformation is applied to its local image coordinates in image  $i$  and image  $j$ . After correcting the distortions, the correspondences are mapped by an affine transformation to a common coordinate system corresponding to the mosaic image. For the set of all images  $B$ , the squared Euclidean distance between correspondence points

should be minimal in the coordinate system of the mosaic image. This goal induces the following optimization problem:

$$\min_{\beta, R, T} \sum_{i,j=1, j \neq i}^B \left( \left\| \left( \phi_d(X^{(i,j)})\beta R^{(i)} + T^{(i)} \right) - \left( \phi_d(X^{(j,i)})\beta R^{(j)} + T^{(j)} \right) \right\|^2 + \lambda \left\| \phi_d(X^{(i,j)})\beta - X^{(i,j)} \right\|^2 \right). \quad (6)$$

The transformation matrix  $R^{(i)}$  and the translation vector  $T^{(i)}$  for each image are obtained from the affine transformation estimated in Eq. (4). The regularization term in Eq. (6) penalizes transformations which map the transformed image points far apart from the original image points. Experiments clearly demonstrate that  $\lambda = 0.01$  is a sufficient weight for the regularization term. In practice the objective can be optimized by iteratively obtaining the affine transformations  $T^{(i)}$  and  $R^{(i)}$  and then the non-linear transformation matrix  $\beta$ . Our experiments demonstrate, that very few iterations are sufficient to obtain  $\beta$ .

When keeping all  $R^{(i)}$  and  $T^{(i)}$  fixed, the solution for  $\beta$  is unique and is obtained by setting the derivative of Eq. (6) with respect to  $\beta$  to zero and then solving for  $\beta$ . This yields

$$\text{vec}(\beta) = \left[ \sum_{i,j=1, j \neq i}^B \sum_{n=1}^{N(i,j)} \left( 2 \cdot R^{(i)} R^{(i)} \otimes \tilde{x}_n^{(i,j)} \tilde{x}_n^{(i,j)} - 2 \cdot R^{(j)} R^{(i)} \otimes \tilde{x}_n^{(j,i)} \tilde{x}_n^{(j,i)} \right. \right. \\ \left. \left. + \lambda (I_{2 \times 2} \otimes \tilde{x}_n^{(i,j)} \tilde{x}_n^{(i,j)}) \right) \right]^{-1} \\ \cdot \text{vec} \left( \sum_{i,j=1, j \neq i}^B \sum_{n=1}^{N(i,j)} \left( -2 \cdot R^{(i)} T^{(i)} \tilde{x}_n^{(i,j)} + 2 \cdot R^{(i)} T^{(j)} \tilde{x}_n^{(j,i)} + \lambda x_n^{(i,j)} \tilde{x}_n^{(i,j)} \right) \right) \quad (7)$$

where  $\tilde{x}^{(i,j)} = \phi_d(x^{(i,j)})$  and  $x_n^{(i,j)}$  denotes the coordinates of the  $n$ th correspondence point between image  $i$  and image  $j$  in the coordinate system of image  $i$ . The index  $n$  of the second sum runs over all  $N(i,j)$  correspondence points between image  $i$  and image  $j$ .

The operator  $\text{vec}(\beta)$  applied to a matrix  $\beta$  concatenates the columns to a vector. The operation  $A \otimes B$  denotes the Kronecker product of an  $m \times n$  matrix  $A$  and an  $r \times q$  matrix  $B$  that yields a  $mr \times nq$  matrix, defined as:

$$A \otimes B = \begin{pmatrix} A_{11}B & A_{12}B & \dots & A_{1n}B \\ A_{21}B & A_{22}B & \dots & A_{2n}B \\ \vdots & & & \vdots \\ A_{m1}B & A_{m2}B & \dots & A_{mn}B \end{pmatrix} \quad (8)$$

#### 4. Evaluation

In this section we experimentally evaluate our calibration approach. Simulated data are employed to test the quality of the correction against known ground truth. In addition we show on real electron microscopy data, that our calibration method is able to reduce distortions with an average uncorrected stitching error larger than 10 pixels to sub-pixel precision. The experiments are designed to focus on different lens systems in order to determine the origin of the distortions in a TEM and to provide guidelines how often a new calibration has to be estimated.

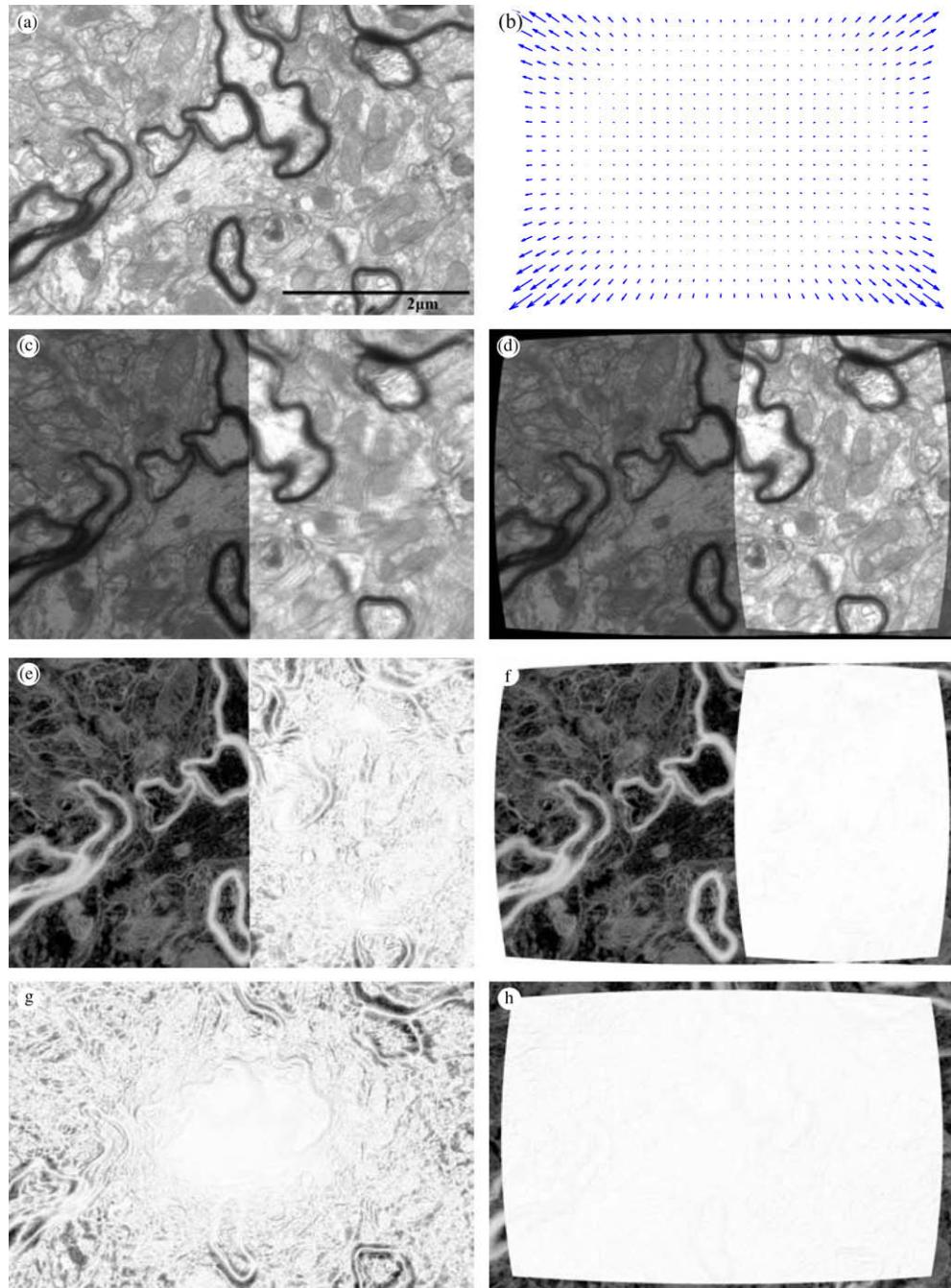
The distortion correction estimation does not require images from special calibration samples. The only limitation is, that the images have enough contrast and texture to find correspondence points with SIFT features and that sufficient image overlap is provided. We propose to take one initial calibration set of nine images

arranged in a  $3 \times 3$  grid with a vertical and horizontal overlap of at least 50% per image in both directions. The large overlap region ensures that all areas of the image, including the center, contribute to the estimated distortion correction. Afterwards the obtained correction transform can be applied to any image taken under equivalent electron optical conditions.

#### 4.1. Simulated ground truth data

To test our calibration method with available ground truth, a set of 9 calibration images ( $1603 \times 1069$  pixels) has been warped by

using the SplineDeformationGenerator ImageJ plugin [Arganda-Carreras et al., 2006](#). Fig. 7 summarizes the results after the images have been unwarped by our method. Subfigure (a) shows one of the nine original images. The distortion applied to the images can be seen in Subfigure (b). To demonstrate the stitching quality we show an image overlay of the example image and another one with 50% overlap. The uncorrected stitching (c) appears blurred in the overlapping area, as the stitching could not seamlessly match image features together. After the correction (d) the images show a satisfactory correspondence. Subfigures (e) and (f) depict the same situation as inverted difference images instead



**Fig. 7.** Evaluation of our calibration approach using ground truth data: (a) ground truth image, (b) the barrel distortion used to distort it; (c) overlay of the distorted ground truth image stitched with its distorted neighbor, (d) stitching after distortion correction; (e) and (f) difference images for the stitchings shown in the row above; (g) difference image of the original image and the distorted version, (h) difference image of the original image and the corrected image. Difference images are shown with inverted contrast to enhance visibility.

of overlays. Here, a dark pixel in the overlapping area demonstrates an error in the stitching. The good quality of the calibration can also be seen by the large reduction of differences between the original and the distorted image (g) when the correction is applied (h). The distortion is removed and the original structure of the image has been recovered. In the distorted images, the average matching error is 6.88 pixels. After the correction had been applied, the error was reduced to 0.41 pixels.

Our method is limited by isotropic scaling since it cannot correct for anisotropic scaling induced by the distortions without knowledge of the original size of image structures. Therefore the regularization term in Eq. (6) ensures that the applied correction changes the scaling factor as little as possible.

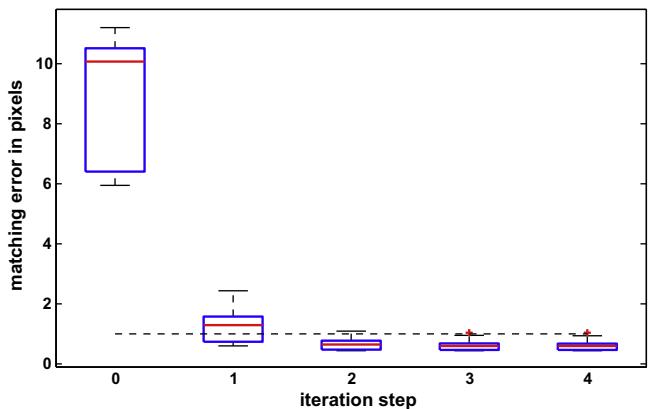
#### 4.2. Electron microscopy data experiments

To test our new method on real world data, we performed a series of experiments. All images ( $4008 \times 2672$  pixels) were acquired with a Philips CM100 equipped with a side-mounted Morada TEM camera from Olympus. The microscope was operated at 100 kV. We used a commercial cross-grating sample, which is a ruled diffraction grating with a periodicity of about 460 nm along two perpendicular axis. The grating is shadowed with gold which yields high contrast in a transmission electron microscope. The grating sample is stable under the electron beam and it is not expected to suffer from severe charging effects. As demonstrated before, our method does not require special calibration samples, but the grid structure of the cross-grating sample enhances the visual impression of the distortion and the correction. The experiments have been designed to answer two questions: (i) The experiments are conducted to measure the reduction of the average stitching error by the distortion correction. (ii) The experiments should allow us to gain a first insight into possible causes and mechanisms of the distortions. In order to distinguish distortions caused by the different lens systems of the TEM, images were captured under the following experimental conditions:

1. condenser lens in over- and under-focus, i.e. cross-over of the beam before and after the specimen,
2. objective lens in focus and in strong under-focus ( $20 \mu\text{m}$ ), i.e. severely different focal lengths,
3. sample not at correct height, compensation for the correspondingly wrong position of the image by the focal length of the objective lens ( $z$ -height of specimen too low, objective at  $20 \mu\text{m}$  over-focus),
4. change of magnification leaving all other parameters unchanged ( $3'400, 19'000, 25'000$  and  $64'000$  times magnification),
5. specimen shifted sideways by some distance, leaving all electron optical parameters unchanged,
6. all settings listed have been checked before and after a full alignment of the microscope.

The following information is expected from these settings: the first setup tests the influence of the condenser lens system; conditions two and three probe the dependency of distortions on the objective lens; test four identifies the influence of the projective system; the setting five is expected to yield information about effects due to the specimen. Procedure six, finally, is performed to measure how sensitively distortions depend on the electron optical parameters. This information is important to judge how often a new distortion correction has to be estimated.

A general analysis over all experiments enables us to measure the performance of the distortion correction. Then the influence of the single parameter settings on the distortion is discussed in relation to the experimental conditions. Fig. 8 shows the results summarized over all experiments in a box plot. The first box, cor-

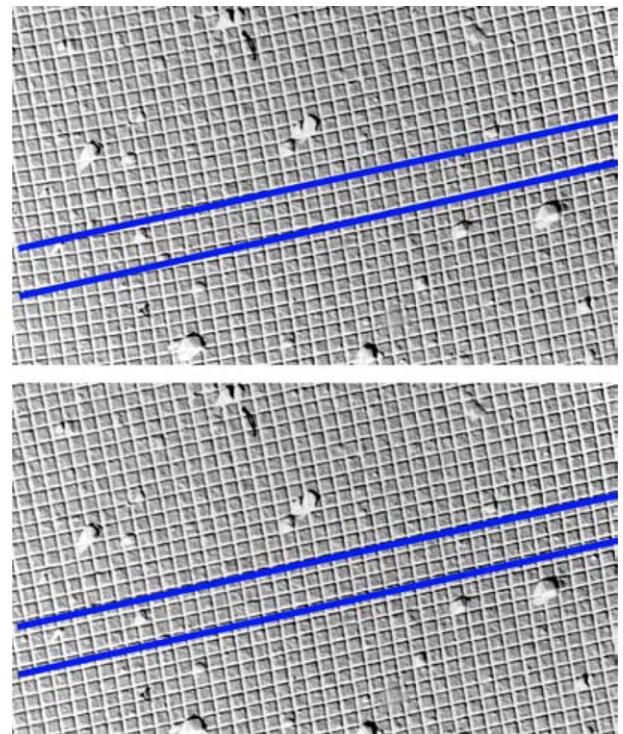


**Fig. 8.** Box plot showing the convergence of the algorithm for 28 different sets of calibration images. On each box, the red line marks the median stitching error in pixels, the edges of the box are the 25th and 75th percentiles, the whiskers extend to the most extreme data points not considered outliers. The dashed black line marks the 1 pixel error boundary. All calibration sets had an initial matching error of more than 10 pixels. As the plot shows, already the second iteration step leads to sub-pixel accuracy. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

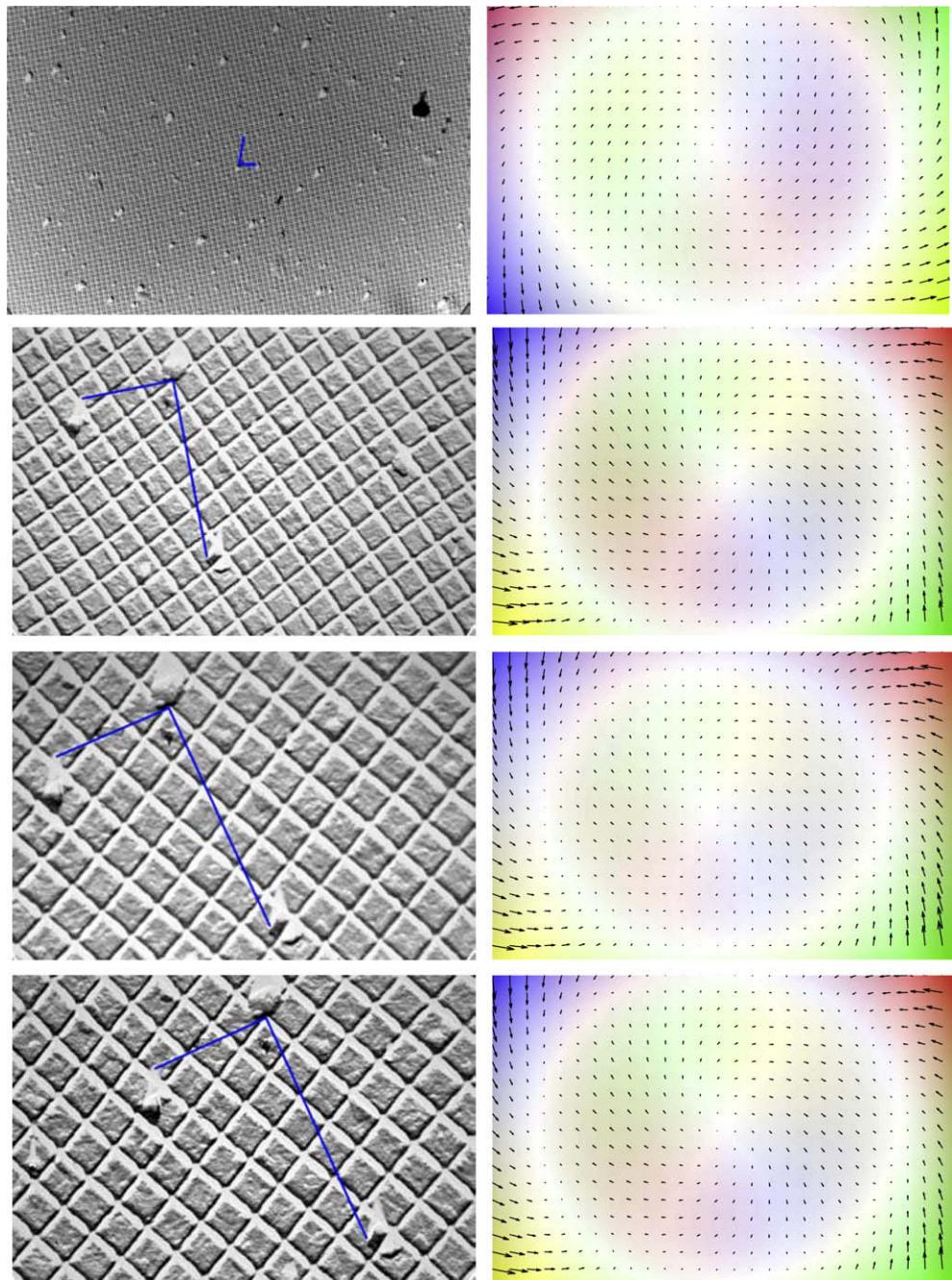
**Table 1**

The table contains the median stitching error (med) and its standard deviation (stddev) over all 28 calibration sets. The error is rapidly reduced with very few iteration steps (iter).

Iter	0	1	2	3	4
Med	10.0708	1.2921	0.6438	0.5968	0.5941
Stddev	1.9843	0.4875	0.1971	0.1637	0.1599



**Fig. 9.** A region of a cross-grating sample image. Top: before correction, bottom: after correction. The distortion correction visibly restores the geometry of the grid structure in the image. In the lower left corner, the distortion is about as large as the grid distance which is 460 nm.



**Fig. 10.** Original distorted images (left) and the estimated distortion correction fields (right) for different magnifications. Arrows are scaled with a factor of three to enhance visibility. In the transformation field images, the hue indicates the direction and the saturation the magnitude of the transformation. Primary magnification from top to bottom: 3'400×, 19'000×, 25'000×, again 25'000× after alignment of the microscope. Between the first two magnifications a clear rotation jump is visible. This is caused by a different setting for the projective lens system. From 19'000× to 25'000× the rotation angle is small. All images are 4008 × 2672 pixels. At the corners the unscaled magnitude of the transformation vectors is larger than 60 pixels.

responding to iteration step 0, depicts the original uncorrected error in the stitching. The large variance of this entry is due to the different parameter settings that cause different distortions. But, already the second iteration step corrects the distortion sufficient enough to reach sub-pixel accuracy. The small variance in the corrected result demonstrates, that the method performs reliably for all parameter settings. The numerical values of the median stitching error and the standard deviation for the single iteration steps are summarized in Table 1.

A visual impression of a distortion and the correction result is depicted in Fig. 9, showing the lower left region of a cross-grating

sample image. When the grid structure is compared to the straight lines drawn in blue, it is clearly visible, that the upper image contains distortions. In the lower left corner the distortions exceed the grid distance of 460 nm. In the bottom image, the correction has been applied and the geometry of the sample is restored.

A comparison of the distortion correction fields, that have been determined by the described procedures, reveals that they are very similar for all settings. The central part of the images shows almost no distortions at all, while severe distortions are observed mainly near the corners of the images (see Fig. 10 for some examples). The stitching error for all single settings is given in Table 2. The most

**Table 2**

The table contains the median stitching error in pixels for all 28 calibration sets. (A) condenser lens in over-focus, (B) condenser lens in under-focus, (C) condenser lens in under-focus and objective in under-focus (20  $\mu\text{m}$ ), (D) condenser lens in under-focus and objective in over-focus (20  $\mu\text{m}$ ) correcting for specimen lower than optimal eucentric height. For magnification settings 3400 and 25,000 experiments were repeated on a second region of the sample.

Mag	3400	3400	19,000	25,000	25,000	64,000
<i>Before alignment</i>						
A	6.24		11.20			
B	6.32	6.92	9.76	10.07	10.65	10.12
C	6.43		11.14	8.43	8.95	
D	6.10	5.95	10.48	10.93	10.30	
<i>After alignment</i>						
A						
B	6.22	6.64		10.69	10.61	9.16
C	6.32			10.47	10.71	
D	6.13		10.22		10.18	

**Table 3**

The table contains the lens currents for the different magnification settings in [mA]. The abbreviations have the following meaning: C1 and C2 are the condenser lenses, OBJ is the objective lens, DIF is the diffraction lens, INT the intermediate lens, P1 and P2 the projective lenses of the Philips CM100.

Mag	3400	19,000	25,000	64,000
C1	391	391	391	391
C2	2601	2601	2599	2601
OBJ	2259	2259	2260	2259
DIF	1098	1619	1613	1579
INT	354	608	657	1007
P1	775	1544	1262	1364
P2	1972	1972	1952	1929

prominent change in the transformation fields can be detected when the magnification changes from 3'400 $\times$  to 19'000 $\times$ . At magnification 3'400 $\times$  the stitching error amounts to six pixels. Varying other microscope parameters or even a full realignment of the electron microscope column has only a marginal effect on the distortion. Increasing the magnification to 19'000 $\times$ , causes the error to grow up to ten pixels for all different parameter settings. The images as well as the estimated transformation fields exhibit a significant jump in rotation (see Fig. 10). This result identifies the projective lens system as the main influencing factor for the distortions.

For different magnifications, the lens currents of the condenser lenses as well as the current of the objective lens remain constant (see Table 3) and thus, the first intermediate image is not changed. A region of this image plane is magnified by the projective lens system. With increasing magnification, the zoom area selected from the center of the intermediate image plane is shrinking. Hence, only the central, little distorted image area is further magnified and imaged by the CCD camera, and a decrease in the distortions for high magnifications would be expected if the condenser and objective lens system would cause the distortions. This analysis identifies the projective lens system as the most likely cause for the distortion change in our experiments. In addition, the increase in the distortions is correlated with a rotation jump, clearly visible in the original images (see Fig. 10). This jump is caused by a different lens configuration for the new magnification setting. As a consequence of this observation we suggest to estimate transformations for distortion correction for each major projective lens configuration. These transformations can then be used as default correction for single images, in cases where no sub-pixel accuracy is required. Table 2 shows, that a careful alignment of the microscope stabilizes the distortions with respect to different focus settings. Hence a good alignment of the microscope increases the benefit of an estimated default distortion

correction field. The largest impact was observed on the under-focus setting, which is widely used to enhance contrast for biological samples.

## 5. Summary and conclusion

In this paper, we present a new method to correct for image distortions caused by electron microscope optics. The method does not require images of prepared calibration samples. Instead it uses redundant information in overlapping image areas to estimate the distortion correction. The method estimates a global transformation as image correction, that applies to all images. Therefore, any structural information in the images is preserved. Our experiments convincingly demonstrate that an arrangement of nine images in a 3  $\times$  3 grid structure with 50% overlap between images defines a reliable configuration to estimate stable distortion fields. We have tested our algorithm on ground truth data, demonstrating that the method is capable of recovering the original image structure. Furthermore, a series of experiments have been conducted to generate real distortions induced by a transmission electron microscope; the distortion correction reduces the average stitching error from over ten pixels to sub-pixel accuracy. The experiments have identified the projective lens system as the main origin for the imaging distortions. This result can be regarded as surprising since so far only the objective lens has been considered for hardware adjustments against spherical and chromatic aberrations. As at the current time adjustments of the projective lens system are not available to the operator of an electron microscope, a software distortion correction like the one presented in this paper is necessary to correct distorted transmission electron microscopy images.

## Acknowledgment

We thank German Köstinger, Nuno Maçarico da Costa, and Kevan Martin from the Institute of Neuroinformatics, UNI-ETH Zurich for providing the TEM images showing biological samples and their support in performing the TEM experiments with the cross-grating sample. Furthermore, we want to thank Roger Wepf, from Electron Microscopy ETH Zurich for valuable discussions about the origin of the distortions and for providing the cryo TEM images. We also thank Stephan Saalfeld of the Max Planck Institute of Molecular Cell Biology and Genetics in Dresden for helpful comments on the distortion model and his support with the Java implementation.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jsb.2010.04.012.

## References

- Akselrod-Ballin, A., Bock, D., Reid, R.C., Warfield, S.K., 2009. Accelerating feature based registration using the johnson-lindenstrauss lemma. MICCAI, 632–639.
- Arganda-Carreras, I., Sorzano, C.O.S., Marabini, R., Carazo, J.M., de Solorzano, C.O., Kybic, J., 2006. Consistent and elastic registration of histological sections using vector-spline regularization. CVAMIA 4241, 85–95.
- Cardona, A., 2006. Trakem2: an imagej-based program for morphological data mining and 3d modeling. In: Proc. First Imagej User and Developer Conference, Luxembourg.
- Claus, D., Fitzgibbon, A.W., 2005. A rational function lens distortion model for general cameras. CVPR 1, 213–219.
- Dauguet, J., Bock, D., Reid, R.C., Warfield, S.K., 2007. Alignment of large image series using cubic b-splines tessellation: application to transmission electron microscopy data. MICCAI, 710–717.
- Devernay, F., Faugeras, O., 2001. Straight lines have to be straight. Mach. Vision Appl. 13 (1), 14–24.
- Fiji, 2010. An image processing package based on imagej (fiji is just imagej – batteries included). URL: <<http://pacific.mpi-cbg.de/>>.

Geyer, C., Daniilidis, K., 2001. Structure and motion from uncalibrated catadioptric views. *CVPR* 1, 279.

Greban, K., Thorpe, C., Kanade, T., 1988. Geometric camera calibration using systems of linear equations. *ICRA*, 562–567.

Hartley, R., Kang, S.B., 2007. Parameter-free radial distortion correction with center of distortion estimation. *IEEE T. Pattern Anal.* 29, 1309–1321.

Kaynig, V., Fischer, B., Wepf, R., Buhmann, J.M., 2007. Fully automatic registration of electron microscopy images with high and low resolution. *Microsc. Microanal.* 13, 198–199.

Koshevov, P., Tasdizen, T., Whitaker, R., Jones, B., Marc, R., 2006. Assembly of large three-dimensional volumes from serial-section transmission electron microscopy. *MIAAB* 10–17.

Lawrence, A., Bouwer, J.C., Perkins, G., Elliman, M.H., 2006. Transform-based backprojection for volume reconstruction of large format electron microscope tilt series. *J. Struct. Biol.* 154 (2), 144–167.

Lowe, D.G., 2004. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vision* 60 (2), 91–110.

Marsh, B.J., 2007. Reconstructing mammalian membrane architecture by large area cellular tomography. *Methods Cell Biol.* 79, 193–220.

Rasband, W.S., 1997–2010. Imagej. Available from: <<http://rsb.info.nih.gov/ij/>>.

Saalfeld, S., Tomančák, P., 2008. Automatic landmark correspondence detection for Imagej. In: Proc. of the Second Imagej User and Developer Conference. Luxembourg.

Sawhney, H.S., Kumar, R., 1999. True multi-image alignment and its application to mosaicing and lens distortion correction. *IEEE T. Pattern Anal.* 21 (3), 235–243.

Shih-Schon, L., Bajcsy, R., 2001. True single view point cone mirror omni-directional catadioptric system. *ICCV*, 102–107.

Stein, G.P., 1997. Lens distortion calibration using point correspondences. *CVPR*, 602–608.

Wang, A., Qiu, T., Shao, L., 2009. A simple method of radial distortion correction with centre of distortion estimation. *J. Math. Imaging Vis.* 35, 165–172.

Zhang, Z., 2000. A flexible new technique for camera calibration. *IEEE T. Pattern Anal.* 22 (11), 1330–1334.