

---

Bioimage informatics

# Image-based correction of continuous and discontinuous non-planar axial distortion in serial section microscopy

Philipp Hanslovsky, John A. Bogovic and Stephan Saalfeld\*

HHMI Janelia Research Campus, Ashburn, VA 20147, USA

\*To whom correspondence should be addressed.

Associate Editor: Robert Murphy

Received on March 25, 2016; revised on December 1, 2016; editorial decision on December 10, 2016; accepted on December 14, 2016

## Abstract

**Motivation:** Serial section microscopy is an established method for detailed anatomy reconstruction of biological specimen. During the last decade, high resolution electron microscopy (EM) of serial sections has become the de-facto standard for reconstruction of neural connectivity at ever increasing scales (EM connectomics). In serial section microscopy, the axial dimension of the volume is sampled by physically removing thin sections from the embedded specimen and subsequently imaging either the block-face or the section series. This process has limited precision leading to inhomogeneous non-planar sampling of the axial dimension of the volume which, in turn, results in distorted image volumes. This includes that section series may be collected and imaged in unknown order.

**Results:** We developed methods to identify and correct these distortions through image-based signal analysis without any additional physical apparatus or measurements. We demonstrate the efficacy of our methods in proof of principle experiments and application to real world problems.

**Availability and Implementation:** We made our work available as libraries for the ImageJ distribution Fiji and for deployment in a high performance parallel computing environment. Our sources are open and available at <http://github.com/saalfeldlab/section-sort>, <http://github.com/saalfeldlab/z-spacing> and <http://github.com/saalfeldlab/z-spacing-spark>.

**Contact:** [saalfelds@janelia.hhmi.org](mailto:saalfelds@janelia.hhmi.org)

**Supplementary information:** [Supplementary data](#) are available at *Bioinformatics* online

---

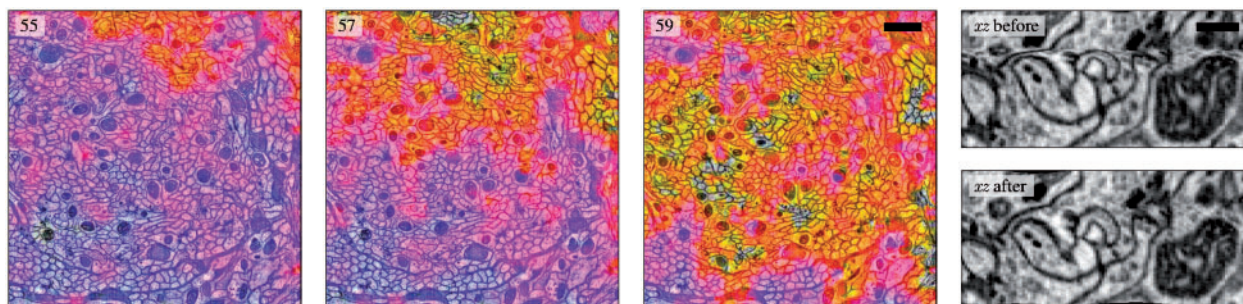
## 1 Introduction

Serial section microscopy has been used for over a century to reconstruct volumetric anatomy of biological samples (Born, 1883). Beyond its classical application in biology, zoology and medical research, serial sectioning in combination with electron microscopy (EM) has become the standard method to reconstruct dense neural connectivity of animal nervous systems at synaptic resolution (Briggman and Bock, 2012; Lichtman *et al.*, 2014; Plaza *et al.*, 2014). A resolution of less than 10 nm per pixel is necessary to separate individual neural processes and to recognize chemical synapses. Sample preparation and data acquisition at this resolution are highly sensitive procedures and, as a result, imaging noise and artifacts during

acquisition can be minimized at best but not entirely avoided. In this paper, we will focus on two major acquisition modalities for large 3-D EM that are used in connectomics: high throughput serial section transmission EM (ssTEM; Bock *et al.*, 2011) and block-face scanning EM with focused ion beam milling (FIB-SEM; Heymann *et al.*, 2006; Knott *et al.*, 2008; Xu and Hess, 2011). While we have developed our methods with a strong focus on these two modalities, we expect them to generalize well to other applications.

### 1.1 Serial section transmission electron microscopy

A series of ultra-thin sections is generated by cutting the plastic embedded specimen using an ultra-microtome with a diamond



**Fig. 1.** Left: Three FIB-SEM  $xy$ -section scans showing *Drosophila melanogaster* neural tissue overlaid with color-coded local  $z$ -spacing, serial index top left. Color overlay was chosen arbitrarily to visualize the wave-like evolution of height variance. Scale bar 1  $\mu\text{m}$ . Right: Magnified crop of an  $xz$ -cross-section of the original (top) and corrected (bottom) series,  $z$ -compression by the “wave” is completely removed. Scale bar 250 nm

knife. Section ribbons are collected on tape (Hayworth *et al.*, 2006) or manually. Manual collection in particular bears the risk of ordering mistakes that in practice occur frequently. The nominal section thickness ranges between 30 and 90 nm, which defines the axial resolution. Yet, discontinuous operation of the ultra-microtome and precision limits of the instrument cause variations of section thickness between and within sections. Shearing forces applied by the knife and during collection introduce deformations to individual sections. Section folds, tears and staining artifacts further complicate the comparison of sections in the series and require that sections be aligned after imaging (Saalfeld *et al.*, 2012). However, compared with block-face SEM as discussed in the next paragraph, ssTEM has two major advantages: (1) sections can be post-stained which results in improved contrast of structures of interest (e.g. synapse T-bars) and (2) imaging is performed in transmission mode which enables high acquisition speed and significantly higher in-plane resolution at high signal to noise ratio.

### 1.2 Focused ion beam scanning electron microscopy

Block-face scanning EM follows a cycle of imaging the block face of a plastic-embedded specimen with a scanning electron microscope (SEM) followed by material removal until complete acquisition of the specimen. In FIB-SEM, FIB milling is used for material removal. This procedure, in practice, generates inhomogeneous  $z$ -spacing and non-planar block faces leading to distorted volumes (Boergens and Denk, 2013; Jones *et al.*, 2014). These distortions exhibit a wave-like evolution of height variances throughout the acquired data and can be severe enough to seriously impede the correct reconstruction of small neural processes (c.f. Fig. 1). FIB-SEM has two major advantages over the previously discussed ssTEM: (1) FIB milling enables significantly higher axial resolution than physical sectioning, which enables the acquisition of isotropic volumes at less than  $(10\text{ nm})^3$  voxel size and (2) fully automatic integration of serial imaging and milling in the vacuum chamber of the microscope bears a lower risk for variances in image quality and provides better initial section alignment and correct section order.

### 1.3 Contribution

Extending our previous work (Hanslovsky *et al.*, 2015), we developed methods for the identification and correction of ordering mistakes as well as planar and non-planar axial distortions through image-based signal analysis without the need for any further apparatus or physical measurements (Section 3). We thoroughly assess efficacy and efficiency in virtual ground truth experiments, demonstrate their applicability to real world problems (Section 4) and compare with state of the art (Supplementary Note A). We

published our methods as open source libraries for the ImageJ distribution Fiji and for deployment in high performance parallel computing environments using Spark (Zaharia *et al.*, 2010).

## 2 Related work

To the best of our knowledge, both post-acquisition order correction for serial section microscopy and non-planar section thickness correction have not yet been addressed in a rigorous way. However, several methods exist for measuring or correcting section thickness or spacing. De Groot (1988) reviews four different methods for estimating section thickness, all of which require additional physical measurements, specialized apparatus, or even destructive modifications of previously acquired sections, rendering the proposed methods impractical or even impossible for certain imaging modalities, e.g. block face SEM. Similarly, Jones *et al.* (2014) introduce an artifact as a fiducial mark from which section thickness can be estimated in FIB-SEM acquisitions under the assumption of planar sections. Berlanga *et al.* (2011) correct small volumes by evening out top and bottom surfaces that have been manually annotated by the user and transforming the whole series accordingly by a single transformation, which fails to capture varying thickness. Boergens and Denk (2013) reduce non-planar distortions during acquisition using measurements of the intensity of the ion beam to control the FIB-SEM milling process. In addition, they estimate section thickness post-acquisition by adjusting  $z$  coordinates such that the peaks of auto-correlations in several  $xz$ -cross-sections have the same half-width in both dimensions.

Most related to our method for image based estimation of planar section thickness is the work by Sporryng *et al.* (2014). They assume an isotropic signal that is sampled at less than isotropic axial resolution, with planar thickness variation. A reference similarity curve is obtained from in-section pixel intensities. The corrected spacing between adjacent sections is then determined by evaluating the inverse of the reference similarity curve at the measured pairwise similarity. We compared Sporryng *et al.*'s method with ours and showed that our global approach performs superior where individual sections are compromised by staining artifacts (Supplementary Note A).

To the best of our knowledge, we are first to propose solely image-based methods for the correction of section ordering mistakes and non-planar axial distortion. Other than existing methods for planar section thickness correction, we do not accumulate pairwise distance estimates relative to a constant reference. Instead, we jointly optimize and update the axial distortion field and an idealized observation along the axial dimension. This allows us to

explicitly account for artifacts that compromise the signal in individual images and to cope with varying properties of the reference signal along the axis of distortion.

### 3 Materials and methods

In the following, we describe our image-based methods for the correction of continuous and discontinuous non-planar axial distortions in serial section microscopy. Our only assumptions are—true for correct section order and spacing—monotonic decrease of pairwise similarity of sections with distance and a slow rate of change of biological tissue, i.e. the shape of the similarity function is locally constant (Section 3.1). Violations of these assumptions indicate wrong section order or spacing. We will describe in detail how coordinate space is transformed to re-establish correctness of these assumptions and thereby correcting section order mistakes (Sections 3.2 and 3.3), planar  $z$ -spacing (Section 3.3) and non-planar  $z$ -spacing (Section 3.4).

#### 3.1 Similarity measure

We define pairwise similarity  $s(P_i, P_j)$  of two sections  $P_i, P_j \in I$ , indexed by their respective positions  $i, j$  along the  $z$ -axis within an image series that has correct section order and spacing, as a symmetric function that decreases strictly monotonically with distance  $|j - i|$ :

$$s : I \times I \rightarrow [0, 1] \subset \mathbb{R} \quad (1)$$

$$(P_i, P_j) \rightarrow s(P_i, P_j) = f(|j - i|) \quad (2)$$

$$|j - i| < |k - l| \Rightarrow f(|j - i|) > f(|k - l|) \quad (3)$$

$$s(P_i, P_j) = s(P_j, P_i) \quad (4)$$

For a series of  $Z$  sections, all pairwise similarities are stored in a  $Z \times Z$  matrix denoted by  $S$  such that

$$S_{ij} = s(P_i, P_j). \quad (5)$$

By definition,  $S$  is a symmetric matrix. In practice, we use noisy surrogate measures for the inaccessible ideal  $s$  such that Equation (3) may not hold for long distances. Thus, for deformation estimation, we ignore measurements for which  $|j - k| > r$ , for a user specified  $r$  that depends on the dataset.

We implemented three similarity measures: (1) the Pearson product-moment correlation coefficient (PMCC) for aligned series, (2) the best block matching coefficient (BBMC) for approximately aligned series and (3) the percentage of true positive feature matches under a transformation model (inlier ratio) for unaligned data. In our experiments (Section 4), we used PMCC and feature inlier ratio.

##### 3.1.1 Pearson product-moment correlation coefficient

The PMCC of two statistical samples  $A, B : |A| = |B| = N$  is defined as

$$\rho_{AB}(A, B) = \frac{\text{cov}(A, B)}{\sqrt{\text{var}(A)\text{var}(B)}} \in [-1, 1] \quad (6)$$

with sample co-variance

$$\text{cov}(A, B) = \frac{1}{N} \sum_{i,j} (A_{ij} - \mu_A)(B_{ij} - \mu_B), \quad (7)$$

where  $\mu_A = \frac{1}{N} \sum_{i,j} A_{ij}$  is the sample mean, and  $\text{var}(A) = \text{cov}(A, A)$  is the sample variance. PMCC is invariant to changes of the mean and

variance of samples  $A$  and  $B$  and therefore robust against contrast and gain variations across the image series. In order to comply with Equation (1), we use

$$s(A, B) = \tilde{\rho}_{AB} = \max(\rho_{AB}, 0) \in [0, 1]. \quad (8)$$

##### 3.1.2 Best block matching coefficient

Similarity estimates using PMCC require the series to be perfectly aligned which, in practice, is not always guaranteed. We therefore implemented an alternative similarity measure that is robust against small local translations, the average over local BBMCs. For any rectangular region  $R_i \subset P_i$ , the best correspondence  $R_j^* \subset P_j$  is determined by maximizing pairwise PMCC over a set of correspondence candidates  $R_j \subset P_j$  of the same width and height, sampled in a small radius around the center of the region. The pairwise similarity of sections  $P_i$  and  $P_j$ ,

$$S_{ij} = \frac{1}{N} \sum_{R_i} \max_{R_j} s(R_i, R_j), \quad (9)$$

is the average of all pairwise similarities between  $R_i$  and corresponding  $R_j^*$ , where  $N$  is the total number of regions  $R_i$  within  $P_i$ .

##### 3.1.3 Inlier ratio

Even BBMC requires that the series is approximately aligned. In ssTEM series, however, approximate alignment is often not available and aligning the series may be impossible because the correct order of sections has not yet been established. To recover the correct order of sections, we need a similarity measure that is independent of alignment. Using transformation invariant features, we match automatically extracted interest points across pairs of sections. We then use a variant of the random sample consensus (RANSAC) in combination with a least squares local trimming estimator (Fischler and Bolles, 1981; Saalfeld *et al.*, 2010) to estimate a model  $M$  that transforms one set of interest points onto the other. The estimator groups all matches into inliers  $\mathcal{I}$  that conform with  $M$  and outliers  $\mathcal{O}$  that do not ( $\mathcal{I} \cap \mathcal{O} = \emptyset$ ). The similarity of two sections is then given by the inlier ratio

$$s(\cdot) = \frac{|\mathcal{I}|}{|\mathcal{I} \cup \mathcal{O}|} \in [0, 1]. \quad (10)$$

For our experiments, we use the scale invariant feature transform (SIFT; Lowe, 2004). Where interest point detection and matching are part of the image alignment pipeline (e.g. Saalfeld *et al.*, 2012), this similarity can be extracted at virtually no cost.

### 3.2 Section order correction

Incorrect section order breaks the monotonicity assumption for similarity measures. With pairwise similarity as a proxy for distance between sections, visiting every section in the correct order is equivalent to visiting every section exactly once on the shortest path possible based on distances derived from pairwise similarity. This can be formulated as an augmented traveling salesman problem (TSP; Applegate *et al.*, 2011; Voigt, 1831). To that end, we represent the image sections  $\{P_i | i = 1, \dots, Z\}$  as vertices  $\mathcal{V} = \{1, \dots, Z\}$  of a fully connected graph  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$  with edges  $\mathcal{E} = \mathcal{V} \times \mathcal{V}$  and associated edge weights  $w(a, b) \forall (a, b) \in \mathcal{E}$ . Based on the intuition that sections are more similar when they are close to one another, we chose  $w(a, b) = 10^5 \times \exp(|1 - S_{ab}|)$  to transfer the similarities into distances (While the monotonicity of this function is equivalent to the simpler  $|1 - S_{ab}|$ , we found that stretching out similarities of

nearby sections helped the TSP solver to find a correct solution). With the addition of a “start” vertex  $\tilde{\mathcal{V}} = \{0\}$  and zero distance edges  $\tilde{\mathcal{E}} = \{(0, i), (i, 0) : \forall i \in \mathcal{V}\}$ ,  $w(a, b) = 0 \forall (a, b) \in \tilde{\mathcal{E}}$ , establishing correct section order is equivalent to solving the TSP for the augmented graph

$$\tilde{\mathcal{G}} = (\mathcal{V} \cup \tilde{\mathcal{V}}, \mathcal{E} \cup \tilde{\mathcal{E}}). \quad (11)$$

The TSP solution will place the start node between the first and the last section of the sorted stack. Assuming that the first section appears before the last section in the initial stack, the correct order can be established by traversing the TSP solution accordingly.

### 3.3 Simultaneous section spacing and order correction

We observed that TSP-sorted series occasionally contain small mistakes such as flipped section pairs. Pairwise comparison alone does not capture global consistency if the similarity measure is too noisy to reliably distinguish between pairs of sections (Fig. 4) because similarity to any neighbor higher than first order is completely ignored. We therefore developed a method that, assuming that sections are in approximately correct order, compares shapes of complete similarity matrices to determine globally consistent order of sections and their relative spacing.

Based on the assumption of monotonically decreasing pairwise similarity and local constancy of the similarity decay, we formulate an optimization problem that simultaneously estimates a real valued position  $\mathbf{c}_i$  for each section  $i$ , a scalar factor  $\mathbf{m}_i$  to compensate for the influence of uncorrelated noise in individual sections to their pairwise similarity scores, and the “true” similarity  $\bar{s}(\cdot)$ . Correct section order can be established by sorting  $\mathbf{c}$  in increasing order. We summarize all variables, parameters and measurements in Table 1.

We forgo any assumptions other than monotonicity and constancy of shape in a local neighborhood. Instead, we estimate the similarity  $\bar{s}_i(d)$  as a function of the distance  $d = k - j$  between two sections  $j$  and  $k$  in a local neighborhood around each section  $i$ . We constrain this neighborhood by a windowing function  $w_s(i, j)$  that specifies the locality of the estimate [Equation (13)]. These local estimates capture both changes in tissue and image properties along the  $z$ -axis. For all  $j$  within this neighborhood, the measured similarities  $S(\mathbf{c}_j, \mathbf{c}_j + d)$  evaluated at integer distances  $d$  from the position of the

**Table 1.** Description of variables and parameters introduced in Equations (13–16)

Input	
$S_{ij}$	Symmetric matrix containing measures of similarity for all pairs of sections indexed by $i$ and $j$ .
Variable	
$i, j, k$	Indices referencing (sub-)sections within the data.
$\mathbf{c}_i$	Location of index $i$ in corrected coordinate space.
$\mathbf{m}_i$	Scaling factor for section $i$ to compensate for independent artifacts.
$S(\cdot)$	$S$ , corrected by $\mathbf{m}$ and warped by $\mathbf{c}$ : $S(\mathbf{c}_i, \mathbf{c}_j) = \mathbf{m}_i \times \mathbf{m}_j \times S_{ij}$
$\bar{s}_i(d)$	Local estimate of the similarity curve around $i$ for all distances $d$ , sampled at integer coordinates, evaluated at $d \in \mathbb{R}$ .
Parameter	
$w_s(i, j)$	Windowing function that specifies the locality of similarity estimates $\bar{s}_i(d)$ .
$w_r(i, j)$	Windowing function to exclude noisy similarity measures of distant sections.

section  $\mathbf{c}_j$  contribute to the similarity function estimate weighted by  $w_s(\cdot)$ . Simultaneously, we warp the coordinate space such that all measured similarities agree with the function estimate [Equation (14)]. In terms of the pairwise similarity matrix that means aligning the contour lines such that they are parallel to the diagonal (c.f. Fig. 2).

Noise in individual sections decreases the pairwise similarity with all other sections in conflict to what  $\bar{s}(\cdot)$  suggests and would thus distort the estimate of  $\mathbf{c}$ . Therefore, we estimate a scaling factor  $\mathbf{m}$  for each section  $i$  to distinguish between displacement and other noise that could distort position correction [Equation (15)]. Using  $\mathbf{m}_i$  and  $\mathbf{m}_j$  to lift all pairwise similarities  $S_{ij}$  closer to  $\bar{s}_i(\mathbf{c}_i - \mathbf{c}_j)$  will account for this effect. Sections that need displacement will not have a consistent bias towards decreased or increased similarities and remain unaffected by this “quality assessment”.

The windowing function  $w_r(\cdot)$  restricts the evaluation of pairwise similarities to a range  $r$  to avoid estimation based on distant sections whose similarity measures tend to be unreliable. In general, we define this window using the Heaviside step function parameterized by range  $r$ ,

$$w_r(i, j) = \theta(r - |j - i|). \quad (12)$$

Each of Equations (13)–(15) contribute to a joint objective [Equation (16)] that is optimized over the function estimate  $\bar{s}(\cdot)$  within the support range constrained by  $w_r(\cdot)$ , the factors  $\mathbf{m}$ , and the coordinates  $\mathbf{c}$ :

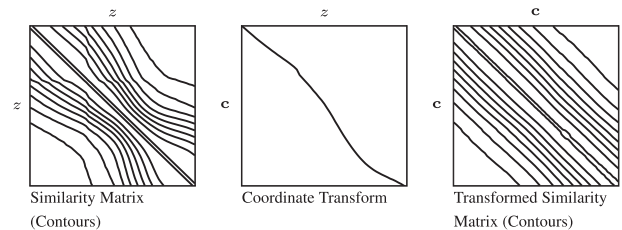
$$\text{SSE}_{\text{fit}} = \sum_i \sum_j w_s(i, j) \sum_k w_r(j, k) (\bar{s}_i(k - j) - S(\mathbf{c}_j, \mathbf{c}_j + k - j))^2, \quad (13)$$

$$\text{SSE}_{\text{shift}} = \sum_i \sum_j w_r(i, j) (\bar{s}_i^{-1}(\mathbf{m}_i \mathbf{m}_j S_{ij}) - (\mathbf{c}_i - \mathbf{c}_j))^2, \quad (14)$$

$$\text{SSE}_{\text{assess}} = \sum_i \sum_j w_r(i, j) (\mathbf{m}_i \mathbf{m}_j S_{ij} - \bar{s}_i(\mathbf{c}_j - \mathbf{c}_i))^2, \quad (15)$$

$$\bar{s}^*, \mathbf{m}^*, \mathbf{c}^* = \arg \min_{\bar{s}, \mathbf{m}, \mathbf{c}} \alpha \text{SSE}_{\text{fit}} + \beta \text{SSE}_{\text{shift}} + \gamma \text{SSE}_{\text{assess}}. \quad (16)$$

We find a local optimum for Equation (16) by alternating least squares. In this optimization scheme, the weights  $\alpha$ ,  $\beta$  and  $\gamma$  do not affect the argmin and are therefore neglected. In the benign case that the series is in approximately correct order and that similarity measures capture sensible information about relative distances between sections, this local optimum is typically the correct solution. We avoid trivial solutions by meaningful regularization: all  $\mathbf{m}_i$  tend toward 1, and  $\mathbf{c}$  is limited by locking the first and last  $z$ -positions. If section order is guaranteed to be correct (as in FIB-SEM), then we do not allow reordering and enforce  $\mathbf{c}_{i+1} - \mathbf{c}_i > 0$  at any iteration.



**Fig. 2.** Warp coordinate space such that contour-lines of transformed similarity matrix  $S(\mathbf{c}_i, \mathbf{c}_j)$  are parallel to diagonal



In addition, we enforce monotonicity of  $\bar{s}$  during estimation of both  $\bar{s}$  and  $\mathbf{c}$ . More precisely, if any scaled similarity estimate  $\mathcal{S}(c_i, c_j)$  violates the monotonicity assumption, this measurement and all subsequent estimates  $\mathcal{S}(c_i, c_k)$  with  $|c_k - c_i| > |c_j - c_i|$  are ignored for this iteration.

### 3.4 Non-planar axial distortion correction

Section spacing estimation (Section 3.3) does not require to consider complete sections but can be applied to any sub-volume defined by a local neighborhood in  $x$  and  $y$  if similarity can be estimated for pairs of sections in that sub-volume. Hence, non-planar deformation fields can be estimated by solving Equation (16) for a grid of independent similarity matrices, each extracted from a local field of view. If grid locations were optimized independently, local smoothness could not be guaranteed which is particularly objectionable as similarity measures typically degrade with a smaller field of view and become increasingly susceptible to noise. Coupling terms between the optimization problems at each grid location would enforce local smoothness but results in a single large optimization problem instead of many independent optimizations. We therefore apply a multi-scale hierarchical approach: Starting with a large field of view—typically the complete section—the spacing between sample points and the field of view around each sample point are decreased at each stage. Both parameters are freely adjustable and can result in overlapping or disjoint grid areas. Local smoothness is enforced through regularization

$$\text{SSE}_{\text{reg}} = \sum_i (c_i - (\lambda \mathbf{b}_i + (1 - \lambda) \bar{c}_i))^2 \quad (17)$$

toward the inferred coordinates  $\mathbf{b}$  at the previous stage. A unique  $\mathbf{b}$  for each sample point at the current stage is generated by interpolated re-sampling of the previous stage. The impact of regularization is controlled by a parameter  $\lambda \in [0, 1]$  with  $\bar{c}$  being the result of Equation (14) at each iteration of the alternating least squares solution of Equation (16). All optimization problems at one stage of the hierarchy depend solely on the results of the previous stage which makes it straightforward to parallelize the solution over all grid cells. The resolution and field of view considered at each stage, the regularization parameters  $\lambda$ , and the range of interest [Equation (12)] for pairwise similarity measurement in the  $z$ -series are exposed as adjustable parameters to the user.

## 4 Experiments

Following the outline of Section 3, we first describe the evaluation of section order correction, both using TSP and section spacing estimation (Section 4.1), before we elaborate on experiments on spacing correction for planar and non-planar distortion (Section 4.2).

### 4.1 Section order correction

We began the evaluation of section order correction with a proof of concept on a small ssTEM dataset (ssTEM-a; ssTEM of *Drosophila melanogaster* CNS, courtesy of D. Bock, R. Fetter, K. Khairy, E. Perlman, C. Robinson, Z. Zheng, HHMI Janelia) of dimensions  $2580 \times 3244 \times 63 \text{ px}^3$  and nominal voxel size  $4 \times 4 \times 40 \text{ nm}^3$ . We perturbed the correctly ordered series using (1) a completely random permutation and (2) a permutation that randomly reassigned the position of sections within a range of  $\pm 4$  for the approaches introduced in Section 3.2 and 3.3, respectively. For (1), we evaluated both PMCC and SIFT inlier ratio as similarity measure from images scaled to 12.5% of their original size. For (2), we used PMCC only.

Similarity matrices before and after section order correction are shown in Figures 3 and 5 for TSP (PMCC and SIFT inlier ratio) and section spacing (PMCC and  $xz$ -cross-section), respectively. Note that, for (2), section order and  $z$ -spacing are estimated simultaneously and therefore the sorted matrix appears warped. TSP re-established the correct section order for both PMCC and SIFT inlier ratio. The run times for optimization of the TSP problems are negligible compared to the time required to extract pairwise similarities (14 ms versus 720 ms for PMCC and 19 ms versus 9064 ms for SIFT inlier ratios). Shorter run times for solving the TSP in the PMCC experiment indicate that, with PMCC, the problem is easier due to better similarity measures. PMCC is superior to SIFT inlier ratio as a similarity measure for well aligned series. Even for larger examples, the run time for the TSP solution remains short, e.g. 2070 ms for 2051 sections (data not shown). All experiments were carried out on a Dell Precision T7610 workstation using the TSP solver *concorde* (Applegate *et al.*, 2006).

With this successful proof of concept at hand, we proceeded with section order correction of a longer section series (ssTEM-b; ssTEM of *Drosophila melanogaster* CNS, courtesy of D. Bock, R. Fetter, K. Khairy, E. Perlman, C. Robinson, Z. Zheng, HHMI Janelia). We chose an unaligned series of 251 complete sections for which we manually curated the correct section order. The objective of the experiment was to re-establish correct section order from an initially unaligned series with ordering mistakes. We therefore extracted the SIFT inlier ratio matrix from the unaligned series and estimated section order via TSP. The solution included small pairwise ordering mistakes. However, these disturbances were sufficiently local to enable elastic alignment (Saalfeld *et al.*, 2012) of the corrected series and to extract a PMCC similarity matrix. We then used the TSP method to estimate order from the PMCC similarities (Fig. 4), decreasing the number of misplaced sections from 2 (0.80%) to zero.

### 4.2 Spacing correction

Similar to the experiments for section order correction, we started with a proof of concept, followed by an extensive experiment for the evaluation of non-planar distortion correction using an artificial ground truth deformation on a real world dataset.

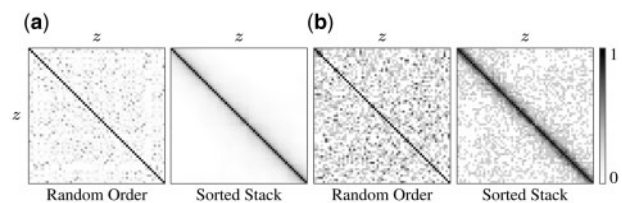


Fig. 3. Similarity matrices for randomly permuted ssTEM-a series before and after section order correction. Similarities were calculated using PMCC (a) and SIFT inlier ratio (b)

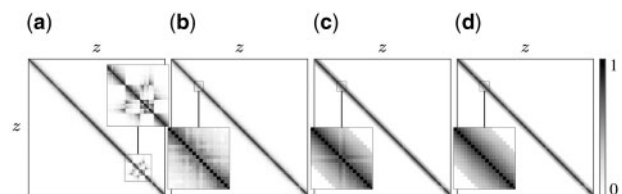
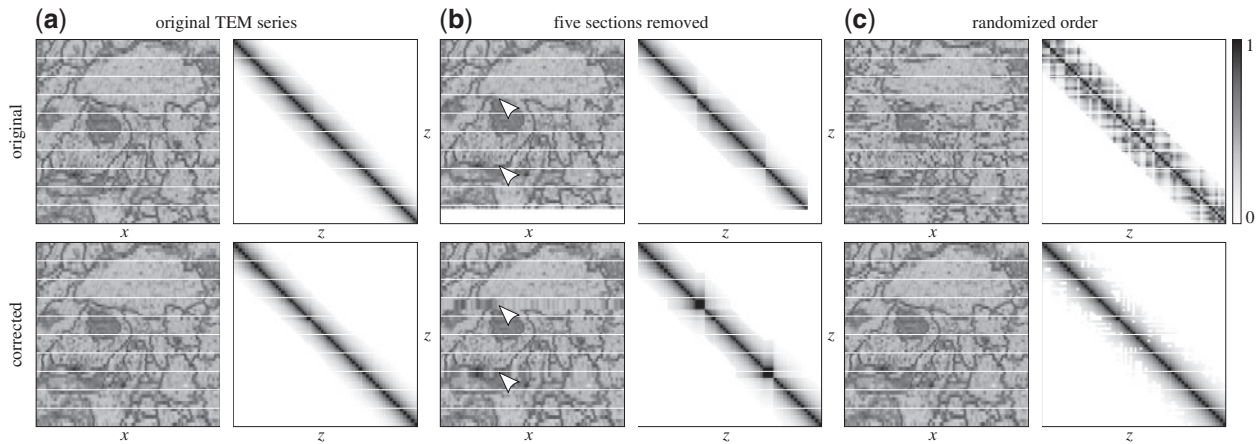


Fig. 4. Section order correction for ssTEM-b. Inlier ratio matrix for original sequence (a) and after correction (b). The major disturbance (bottom right) could be resolved but two sections remain flipped (magnified view). This becomes more apparent in the PMCC matrix of the aligned series (c). Repeated TSP correction resolves this remaining issue (d)



**Fig. 5.** z-position correction experiments for ssTEM-a: original series (a), missing sections (b), and randomized order (c) with a shared coordinate frame in z, as indicated by the white grid. Top/bottom show an xz-cross-section (left sub-column) and corresponding intensity-encoded pairwise similarity matrices (right sub-column) before/after z-position correction. Arrows in the center column highlight removed sections

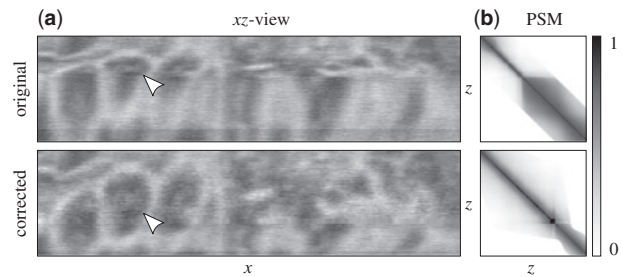
#### 4.2.1 Section spacing correction

For the evaluation of section spacing correction, we corrected and visually inspected distortions in two datasets: ssTEM-a and FIB-SEM-a (FIB-SEM of *Drosophila melanogaster* CNS, courtesy of K. Hayworth, H. Hess, C. Shan Xu, HHMI Janelia, Hayworth *et al.*, 2015) with dimensions  $2048 \times 128 \times 1000 \text{ px}^3$  and nominal voxel size  $8 \times 8 \times 2 \text{ nm}^3$ . The latter is an excerpt of a larger dataset with dimensions chosen such that axial distortions can be considered approximately planar.

Figure 5 shows xz-cross-sections and the according similarity matrices before and after section spacing correction for (a) the original ssTEM-a dataset, (b) Sections 20, 21, 22, 46, 48 removed and (c) randomized section order. Our experiments show that for the original dataset, z-spacing varies between 0.6 and 1.6 px (24 and 64 nm). Section spacing correction of (b) and (c) was evaluated by comparing the estimated transformations with the result of (a) as “ground truth”. The estimated transformation for (b) correctly stretches the data where sections were removed and deviates (absolute value) from the ground truth by 0.13 px (5.2 nm) on average, and not more than 0.28 px (11.2 nm). Sections removed for this experiment do not contribute to the evaluation. For the simultaneous order and spacing correction (c), we measured an absolute deviation from the ground truth of 0.044 px (1.76 nm) on average, and not more than 0.13 px (5.2 nm). All ssTEM-a section spacing correction experiments finished in 0.6 s (similarity matrix calculation) and 0.4 s (inference, 100 iterations) on a Dell Precision T7610 workstation using the default parameters of the provided Fiji plugin.

We observed stronger distortions in FIB-SEM-a as shown in Figure 6 (top). Stretched/condensed regions are highlighted in an xz-cross-section and appear in the respective similarity matrix as regions with slow/fast decay of similarity. After section spacing correction (Fig. 6 bottom), the corrected xz-cross-section appears homogeneously sampled and similarity decay is approximately constant. The estimated section spacing varies between 0.14 and 10.2 px or 0.28 and 20.4 nm. On the Dell Precision T7610 workstation used for this experiment, similarity matrix estimation and inference (150 iterations) took 62.3 and 49.4 s, respectively, using the default parameters of the provided Fiji plugin, with the exception of  $r = 55$  [Equation (12)].

On artificial distortions of dataset ssTEM-a, we compared our method with the approach by Spurring *et al.* (2014) (Supplementary



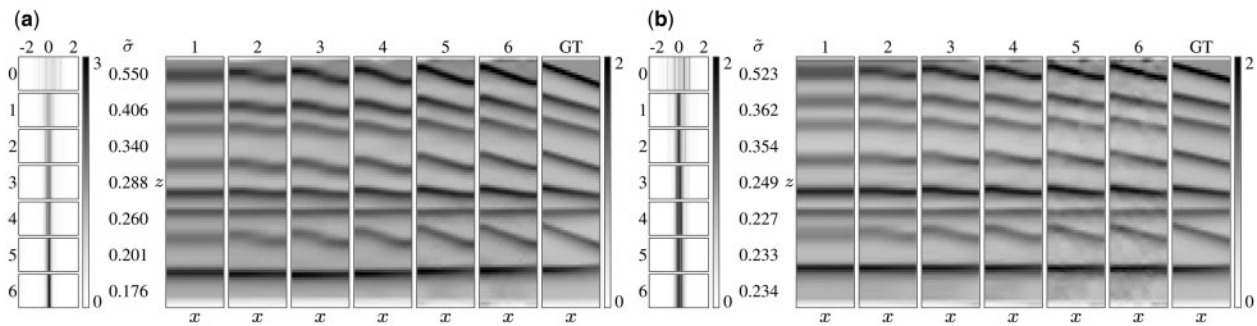
**Fig. 6.** z-position correction experiment for FIB-SEM-a: Top/bottom show an xz-cross-section (a) and corresponding intensity-encoded pairwise similarity matrices (PSM; b) before/after z-position correction. (a) and (b) share the same coordinate frame in z. Arrows highlight areas that are visually stretched or compressed in the original acquisition and appear biologically plausible after correction

Note A) demonstrating that our method performs superior if individual sections are compromised by staining artifacts.

#### 4.2.2 Non-planar distortion correction

We evaluated the performance of non-planar deformation correction against synthetic ground truth. To that end, we applied synthetic non-planar axial distortion to a distortion free reference series, estimated the distortion with our method (Section 3.4) and compared the estimate with the synthetic ground truth. Since distortion free volumes do not exist, we had to first correct the original image volume using the same non-planar axial distortion correction method. The resulting series, from the perspective of our method, is free of distortions. To compensate for the apparent bias in this approach, we ran our experiment not only in the original orientation but permute the coordinate axes such that the new axial dimension falls into the unprocessed image plane.

The data used in this experiment are a subset of FIB-SEM-b (FIB-SEM of *Drosophila melanogaster* CNS, courtesy of K. Hayworth, H. Hess, C. Shan Xu, HHMI Janelia, Hayworth *et al.*, 2015) with dimensions  $4000 \times 2500 \times 2100 \text{ px}^3$  and voxel resolution  $8 \times 8 \times 2 \text{ nm}^3$ . Initial non-planar axial distortion correction was distributed onto 60 compute nodes with 16 cores each and took 120 minutes to finish. We scaled the corrected series along the z-axis by a factor of 0.25 resulting in an isotropic volume of  $4000 \times 2500 \times 525 \text{ px}^3$  from which we extracted two sub-volumes: (a) 100



**Fig. 7.** Normalized histograms (left) of differences between estimated transformation and ground truth (GT) and visualization of the estimated transformation for all stages and GT for experiments on both sub-volumes (a) and (b) via  $xz$  cross-sections of the gradient (right). Histogram bins range from -2 px to 2 px, with maximum counts of 3 and 2 for (a) and (b), respectively. The gradients range from 0 px to 2 px

complete  $xy$ -sections starting at  $z = 25$ , and (b) 100  $xz$ -cross-sections of dimension  $3000 \times 475$  px<sup>2</sup> starting at  $y = 1000$ . For (b), we flipped the  $y$ - and  $z$ -axes such that the synthetic distortion could be consistently applied along the  $z$ -axis. Our synthetic distortion model is this: Randomly oriented planes superimposed with trigonometric functions act as attractors that shift the coordinates towards the attractor along the  $z$ -axis as a monotonically decreasing function of the distance to the attractor along  $z$ . This generates waves and plateaus that approximately resemble phenomena that we observed in the original volume before pre-correction. We then applied non-planar distortion correction to the synthetically deformed series as described in Section 3.4 (parameters listed in [Supplementary Note B](#)). We compared estimated and ground truth distortions at every stage of the hierarchical solution and show histograms of the pixel-wise differences (Fig. 7). Since we are not interested in low frequency distortion of the volume, we mapped each estimate onto the ground truth using a linear transformation that minimizes the squared difference of corresponding look-up table entries within local support defined by a Gaussian window with  $\sigma = (\sigma_x, \sigma_y, \sigma_z)$ .

The evolution of the estimated distortion for each of the sub-volumes is shown in Figure 7. For intuitive visualization, the gradient is displayed. Starting at a complete field of view and a resolution of 1 px<sup>2</sup> in  $x$  and  $y$  at stage 1 (planar estimate), the field of view/resolution is decreased/increased by a factor of two in both  $x$  and  $y$  with every sub-sequential stage which allows for a more accurate estimate of the deformation. At the same time, noise in the data will have a stronger influence on smaller fields of view (c.f. Fig. 7, stage 6) and sets a limit to the resolution at which the deformation can be estimated. The histograms of differences did not improve after (a) stage 6 or (b) stage 4. We chose  $\sigma = (\infty, \infty, 120$  px) for the Gaussian window to estimate the linear transformation. The mean of differences between estimate and ground truth is approximately zero for all stages. We therefore used the standard deviation of the error  $\bar{\sigma}_i$  for stage  $i$  including the baseline  $i = 0$  as a quality indicator. Smaller  $\bar{\sigma}_i$  means better estimates of the ground truth. For (a), we found  $\bar{\sigma}_0 = 0.550$  px and  $\bar{\sigma}_6 = 0.176$  px, and for (b), we found  $\bar{\sigma}_0 = 0.523$  px and  $\bar{\sigma}_4 = 0.227$  px. As expected, non-planar axial distortion correction considerably decreased the distortion of the series in both experiments.

## 5 Discussion

We developed novel methods to address two previously unsolved problems: (1) establish the correct order of unordered section series and (2) compensate for planar and non-planar axial distortion. We demonstrated through extensive experiments that our methods

work reliably and with high accuracy and efficiency on both ssTEM and FIB-SEM data. We went beyond pure proof of concept and showed that our methods are applicable to and perform well on large real world datasets.

In large ssTEM series, the combination of automatic alignment and series sorting has the potential to greatly reduce the need for manual intervention. Non-planar axial distortion correction addresses the peculiar wave-problem in FIB-SEM which, we believe, will have a strong impact on the future application of FIB-SEM for high resolution 3-D reconstruction.

In this work, we made only mild assumptions about the data, i.e. monotonic decrease of pairwise similarity and local constancy of the shape of the similarity curve. While this means that our methods can be applied to a wide range of data, we predict that many problems would benefit from domain specific modeling. For example, explicit modeling of FIB-SEM-waves has the potential to further increase the accuracy of the estimated deformation field. We will work on these ideas in our future research.

## Acknowledgements

This work was supported by HHMI. We thank Davi Bock, Ken Hayworth, Harald Hess, Richard Fetter, Khaled Khairy, Eric Perlman, Camenzind Robinson, Shan Xu, and Zhihao Zheng for data and valuable discussion.

*Conflict of Interest:* none declared.

## References

- Applegate, D. *et al.* (2006) *Concorde TSP solver*. William Street, Princeton, NJ.
- Applegate, D.L. *et al.* (2011) *The Traveling Salesman Problem: A Computational Study*. Princeton University Press.
- Berlanga, M.L. *et al.* (2011) Three-dimensional reconstruction of serial mouse brain sections: Solution for flattening high-resolution large-scale mosaics. *Front Neuroanatomy*, 5, 17.
- Bock, D.D. *et al.* (2011) Network anatomy and in vivo physiology of visual cortical neurons. *Nature*, 471, 177–182.
- Boergens, K.M. and Denk, W. (2013) Controlling FIB-SBEM slice thickness by monitoring the transmitted ion beam. *J. Microsc.*, 252, 258–262.
- Born, G.J. (1883) Die Plattenmodellirmethode. *Archiv Für Mikroskopische Anatomie*, 22, 584–599.
- Briggman, K.L. and Bock, D.D. (2012) Volume electron microscopy for neuronal circuit reconstruction. *Curr. Opin. Neurobiol.*, 22, 154–161.
- De Groot, D.M. (1988) Comparison of methods for the estimation of the thickness of ultrathin tissue sections. *J. Microsc.*, 151, 23–42.

- Fischler, M.A. and Bolles, R.C. (1981) Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM*, **24**, 381–395.
- Hanslovsky, P. et al. (2015) Post-acquisition image based compensation for thickness variation in microscopy section series. In: *ISBI'15*, pp. 507–511.
- Hayworth, K.J. et al. (2006) Automating the collection of ultrathin serial sections for large volume TEM reconstructions. *Microsc. Microanal.*, **12**, 86–87.
- Hayworth, K.J. et al. (2015) Ultrastructurally smooth thick partitioning and volume stitching for large-scale connectomics. *Nature Methods*, **12**, 319–322.
- Heymann, J.A.W. et al. (2006) Site-specific 3D imaging of cells and tissues with a dual beam microscope. *J. Struct. Biol.*, **155**, 63–73.
- Jones, H.G. et al. (2014) Investigation of slice thickness and shape milled by a focused ion beam for three-dimensional reconstruction of microstructures. *Ultramicroscopy*, **139**, 20–28.
- Knott, G. et al. (2008) Serial section scanning electron microscopy of adult brain tissue using focused ion beam milling. *J. Neurosci.*, **28**, 2959–2964.
- Lichtman, J.W. et al. (2014) The big data challenges of connectomics. *Nature Neurosci.*, **17**, 1448–1454.
- Lowe, D.G. (2004) Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.*, **60**, 91–110.
- Saalfeld, S. et al. (2010) As-rigid-as-possible mosaicking and serial section registration of large ssTEM datasets. *Bioinformatics*, **26**, i57–i63.
- Saalfeld, S. et al. (2012) Elastic volume reconstruction from series of ultra-thin microscopy sections. *Nature Methods*, **9**, 717–720.
- Sporring, J. et al. (2014) Estimating the thickness of ultra thin sections for electron microscopy by image statistics. In: *ISBI'14*, pp. 157–160.
- Plaza, S.M. et al. (2014) Toward large-scale connectome reconstructions. *Curr. Opin. Neurobiol.*, **25**, 201–210.
- Voigt, B.F. (1831) Der Handlungsreisende, wie er sein soll und was er zu thun hat, um Aufträge zu erhalten und eines glücklichen Erfolgs in seinen Geschäften gewiss zu sein. *Commis-Voageur, Ilmenau*.
- Xu, C. and Hess, H. (2011) A closer look at the brain in 3D using FIB-SEM. *Microsc. Microanal.*, **17**, 664–665.
- Zaharia, M. et al. (2010) Spark: Cluster computing with working sets. In: *Proceedings of the 2nd USENIX Conference on Hot Topics in Cloud Computing*, HotCloud'10, pp. 10–10. USENIX Association, Berkeley, CA, USA.