Elastic volume reconstruction from series of ultra-thin microscopy sections

Stephan Saalfeld[1], Richard Fetter[2], Albert Cardona[3,2] & Pavel Tomancak[1]

Anatomy of large biological specimens is often reconstructed from serially sectioned volumes imaged by high-resolution microscopy. We developed a method to reassemble a continuous volume from such large section series that explicitly minimizes artificial deformation by applying a global elastic constraint. We demonstrate our method on a series of transmission electron microscopy sections covering the entire 558-cell *Caenorhabditis elegans* embryo and a segment of the *Drosophila melanogaster* larval ventral nerve cord.

Serial-section microscopy is a classic technique for detailed anatomical reconstruction of large biological specimens. Typically, the fixed specimen is embedded in a block of solid medium and then cut into a series of ultra-thin sections. Sections are collected, mounted, individually stained and imaged. Using ultra-thin sections effectively eliminates the penetration problem for both staining and imaging. Furthermore, the minimum achievable section thickness at less than 40 nm is a significant improvement over the axial resolution that can be achieved by optical sectioning techniques such as confocal laser scanning microscopy. Sections can be imaged as mosaics of overlapping image tiles, either manually or using a motorized stage, which allows for the imaging of large fields of view. In combination, these advantages render serial-section microscopy particularly useful for large-scale high-resolution reconstructions of dense neuronal tissue, where the method, mediated by electron microscopy (EM), recently experienced a renaissance[1–6].

The downside of the method is that physically cutting a block into sections destroys the continuity between sections and leads to deformation of individual sections. To recover the imaged volume and extract biologically interesting information, as with the reconstruction of neuronal circuits[2,3,5], sections need to be aligned and distortion must be removed. Alignment can be achieved by maximizing the overlap of similar image content between adjacent sections. However, there are two unknowns that change image content across the section series: specimen shape and independent section distortion introduced during preparation. Naively warping one section into another would compensate for the shape of the specimen and introduce artificial deformation. Our method exploits the fact that the biological specimen's shape typically changes smoothly across sections, whereas the independent distortion in each section is random and uncorrelated with neighboring sections. We align all sections not only to their direct neighbors in the series but to all sections in a local neighborhood by modeling sections as two-dimensional (2D) elastic sheets that penalize nonrigid deformation (**Fig. 1a** and **Supplementary Fig. 1**). All sections are treated as moving targets in a template-free global alignment. The elastic constraint is implemented as a spring-connected particle system in which each section is represented as a triangular spring mesh (**Fig. 1b**, **Supplementary Video 1** and Online Methods).

For each vertex of the spring mesh, we search for the corresponding location in other sections by pairwise block matching using normalized cross-correlation (NCC). To that end, we explore all translation vectors in an immediate neighborhood, which requires sections to be in approximate alignment (**Fig. 1c,d**). We estimate this approximate alignment using automatically extracted landmark correspondences from invariant local image features as described previously[7,8]. Originally proposed for robust object recognition under partial occlusion, this method can deal with significant nonlinear distortion and image artifacts that inevitably occur in large section series (**Supplementary Fig. 2**). Matching local image features and matching local blocks both create a substantial number of spurious matches that would impair alignment and introduce artificial deformation. We effectively remove such spurious matches using a set of filters that include local properties of the features and the block matches as well as global geometric constraints imposed by the supported transformation (**Fig. 1e,f**, **Supplementary Fig. 3** and Online Methods). The ratio of matches passing the filters constitutes a deformation-invariant similarity metric for two sections that can be used to correct the order of the series or to estimate the number of missing sections (**Supplementary Fig. 4**).

All vertices for which corresponding locations in other sections could be identified are connected by zero-length springs to those sections (**Fig. 1a,b**). The distance in a series to which cross-section connections spread is limited by how rapidly the biological structure changes across sections (for ~50-nm transmission electron microscopy section series (ssTEM), it is typically 7 ± 5 sections). Springs across and within sections serve concurrent purposes: whereas cross-section connections support series alignment, springs in the triangle mesh within sections

Figure 1 | The elastic alignment method. (**a**) All sections in the series are aligned not only to their direct neighbors but to all sections in a local neighborhood. Sections are shaded to visualize how the influence of cross-section connections decreases in inverse proportion to the distance between the two sections in the series. That influence is specified by the spring constant. (**b**) Sections are modeled as elastic sheets by a 2D spring-connected triangle mesh. Springs within the mesh stabilize the section. Springs across sections are depicted by orange arrows; they have a relaxed length of 0 and drag the sections toward alignment.

(**c**) Corresponding landmarks in two adjacent electron microscopy sections that were established using local invariant features are connected by lines. (**d**) These landmarks are used to calculate an initial approximate alignment, and the remaining local deformation is estimated by block matching, visualized here by lines connecting the corresponding locations. (**e**) The resulting deformation field is displayed as color-intensity–encoded displacement vectors. Orientation-length scale (small circle) is magnified for better visualization. (**f**) Spurious matches show up as outlier colors and are automatically rejected using local and global filters.
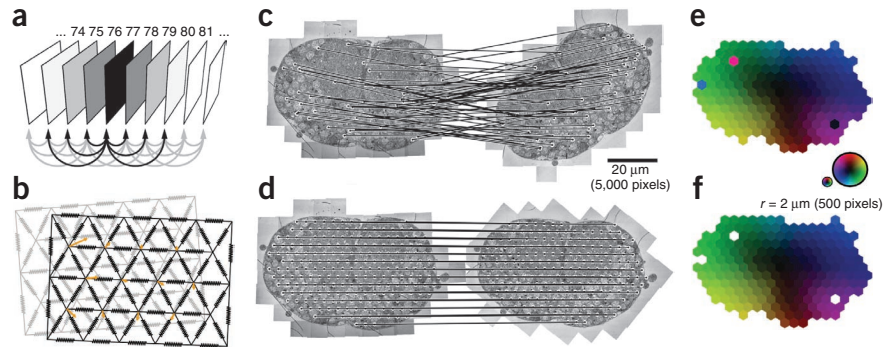
tend toward maintaining a rigid transformation of the sections and penalize distortion. Relaxing this system leads to a series of smoothly aligned sections with the required nonrigid deformation distributed equally among all sections. That is, for each individual section, the deformation relative to a rigid transformation is explicitly minimized (**Supplementary Fig. 5**). Because of this constraint, arbitrarily large section series can be aligned without propagating transformation errors. Similarly to elastically aligning a series of deformed serial sections, our method can be used to assemble montages from deformed overlapping image tiles covering a single section (**Supplementary Fig. 6**). Taken together, a single framework enables the montaging and alignment of massive series of tiled sections.

Similar elastic constraints have been proposed earlier[9,10] that combine a search for an elastic alignment and a pixel-based pairwise similarity estimate between adjacent sections in an iterative solution. This previous work proposes initial linear prealignment of the section series based on variants of principal component analysis. Our method differs in four key areas: first, we compare and align not only adjacent sections but all sections in a local neighborhood (**Fig. 1a** and **Supplementary Fig. 1**). Second, we use invariant local image features[7] to calculate an initial approximate alignment[8] (**Fig. 1c**). Third, we separate the pairwise correspondence search from the elastic alignment, yielding an efficient solution for even very large data (**Fig. 1d**). Fourth, we implement a set of filters to robustly exclude staining artifacts and otherwise corrupted image regions from contributing to the alignment (**Fig. 1e** and **Supplementary Fig. 3**). To evaluate our method quantitatively, we generated synthetic volumes that mimic the properties of biological tissue, sectioned them and introduced artificial deformation to the sections. We measured the alignment error using a sample of straight lines projected through the volume along the z axis. The elastic method outperformed rigid and affine alignment in its ability to recover the straight lines (**Supplementary Figs. 7–12** and **Supplementary Videos 2** and **3**).

We applied our method to two large ssTEM data sets (**Supplementary Table 1**) using a standard quad-core desktop computer with 24 gigabytes of memory. The first data set is a series showing an entire threefold stage *C. elegans* embryo. We scanned 803 sections of 50 nm thickness from film
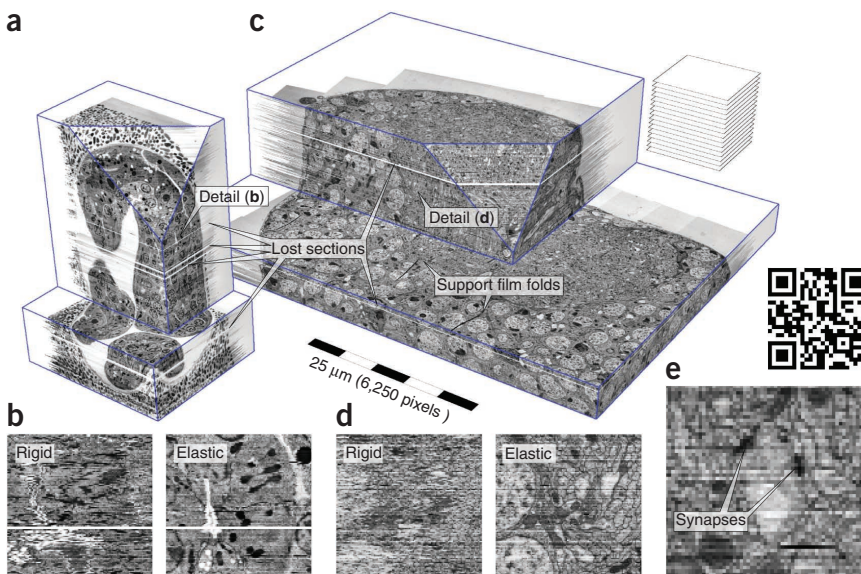

Figure 2 | Reconstruction of two exemplary TEM section series. (**a**–**e**) Sections were scanned from film negatives (**a**,**b**) or assembled from many overlapping digital camera images (**c**–**e**) using our elastic alignment method in montaging mode. Parts of the reconstructed volumes are shown as arbitrarily sliced 3D renderings (**a**,**c**). The planar resolution (scale bar, ~4 nm per pixel) is ~10× higher than the axial resolution (40–50 nm per section). The orientation of the section series is orthogonal to the horizontal plane (see stack, right). Specimens shown are a threefold *C. elegans* embryo (803 sections; **a**,**b**) and 1.5 segments of the ventral nerve cord of a first-instar *Drosophila* larva (458 sections, each section consists of ~70 overlapping image tiles; **c**–**e**). (**e**) Image showing individual synapses in the orthogonally re-sliced volume. The Quick Response (QR)-code links to a collection of videos at http://fly.mpi-cbg. de/elastic/.

Figure 3 | Comparison of the reconstructed shapes of neuronal arbors using rigid series alignment and our elastic method. Exemplary neuronal arbor skeletons were manually traced in the *Drosophila* series using the TrakEM2 software. The resulting shapes are compared for elastic (**a**,**b**) and rigid (**c**,**d**) series alignment. Traces are shown in two perspective projections: dorsal view (**a**,**c**) and lateral view from left to right (**b**,**d**). The section plane is orthogonal to the projection plane; therefore, longitudinal branches expose jitter where alignment insufficiently compensates for low-scale distortion (arrowheads and inset). Asterisks indicate a noticeable misalignment due to a gap of five lost sections (inset). Note that in the rigidly aligned series (**c**,**d**), this misalignment cannot be distinguished from general jitter.
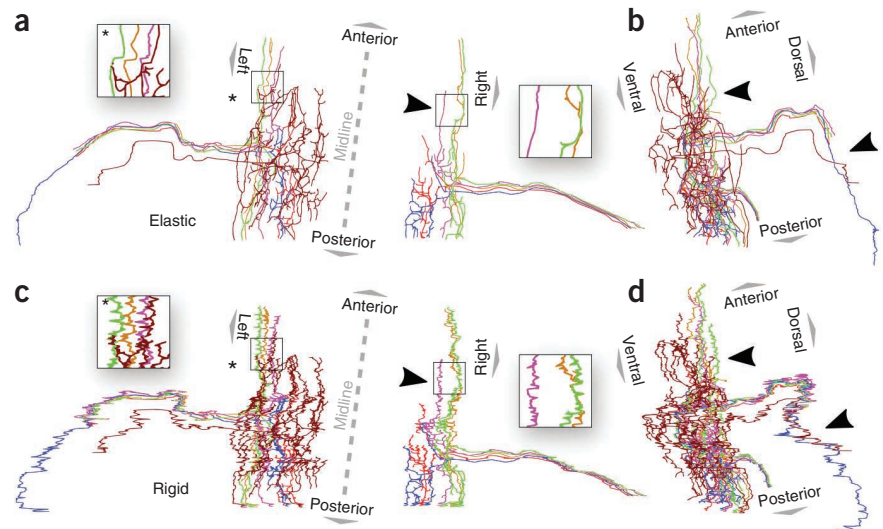
negatives at a size of 6,160 × 4,640 pixels, which resulted in a resolution of 4 nm per pixel (**Fig. 2a**). The series was prealigned rigidly and then aligned elastically by exploring a neighborhood of up to six sections for each section. The elastic method dramatically improved the alignment both in terms of overall specimen outer shape and the internal structure (**Fig. 2b** and **Supplementary Videos 4–7**). We made the result available for interactive exploration at various scales on the data sharing platform CATMAID[11] (http://fly.mpi-cbg.de/c-elegans).

The second data set is an approximately transversal series through an abdominal segment of the ventral nerve cord of a *D. melanogaster* first-instar larva (**Supplementary Fig. 13**). The series consists of 458 sections at 45 nm thickness, each imaged as a mosaic of more than 70 overlapping image tiles (33,051 images all together) of 2,048 × 2,048 pixels covering a canvas of about 22,000 × 17,000 pixels at a resolution of 4 nm per pixel (**Fig. 2c**). Transmission electron microscope (TEM) sections experience heat-induced deformation during image acquisition, which resulted in displacements of up to 50 pixels when only a rigid transformation was used to stitch the montages (**Supplementary Fig. 6**). Consequently, this data set was aligned in two elastic alignment steps: first, all sections were elastically montaged and second, the series of montages was elastically aligned by exploring a neighborhood of up to 8 sections. In contrast to the procedure used with the *C. elegans* data set, we initialized the *Drosophila* elastic series alignment with the result of a previously developed automatic landmark-based method[8] (instead of a rigid alignment). As with the *C. elegans* data set, we observed dramatic improvement of the alignment after the elastic method was applied to the *Drosophila* data set, in terms of both the ventral nerve cord's outer shape and the internal structure down to the resolution sufficient for distinguishing individual synapses in the axial direction (**Fig. 2d**,**e**, **Supplementary Fig. 14** and **Supplementary Videos 8–11**).

To further substantiate the benefits of our elastic alignment method for recovering the biological shape of the imaged specimen, we traced several individual neurons from their cell bodies to the neuropil where they branch and engage in synaptic connections. Tracing was performed manually, using the TrakEM2 software[12], on a previous version of the data set aligned using manually corrected sequential affine transformations comparable in quality to the rigid alignment shown in this manuscript. The manual traces were computationally transferred into the elastically aligned data set and visualized (**Fig. 3**). Whereas rigid alignment suffers from characteristic jitter of traced neuronal profiles, the elastically aligned data set is smooth and better reflects shape details of the biological tissue. Jitter from insufficient alignment contributes notably to the total length of skeleton traces. Elastic alignment reduces the total skeleton length of the neuronal arbors shown in **Figure 3** from 2.87 mm in the rigidly aligned series and 1.55 mm using our previous method[8] to 1.25 mm, which approaches the lower bound length of the skeleton graph of 0.95 mm (**Supplementary Figs. 15** and **16**). The ability to extract better axonal shapes will aid in the comparison of EM and light microscopy data for neuronal circuit reconstruction at vastly different scales[3,13].

We have implemented our elastic alignment method in the Java programming language on top of the popular image processing program ImageJ. The method is available through two standalone ImageJ plug-ins (for creating montages and series alignment) and embedded in the registration and annotation toolkit TrakEM2, where it is complemented by other registration, segmentation and data mining tools[12]. The method is released as open source under the General Public License and distributed through the ImageJ distribution Fiji[14] (**Supplementary Note**). In principle it can be applied to reconstruct any large serial-section data set such as array tomography[15] (**Supplementary Videos 12–14** and **Supplementary Note**). These properties make this method ideally placed for application to emerging and future challenges in high-resolution reconstruction of large biological specimens imaged as series of physical sections.

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Supplementary information is available in the online version of the paper.

1. Hayworth, K.J., Kasthuri, N., Schalek, R. & Lichtman, J.W. *Microsc. Microanal.* **12** (suppl. 02), 86–87 (2006).
2. Anderson, J.R. *et al. PLoS Biol.* **7**, e1000074 (2009).
3. Cardona, A. *et al. PLoS Biol.* **8**, e1000502 (2010).
4. Chklovskii, D.B., Vitaladevuni, S. & Scheffer, L.K. *Curr. Opin. Neurobiol.* **20**, 667–675 (2010).
5. Bock, D.D. *et al. Nature* **471**, 177–182 (2011).
6. Briggman, K.L. & Bock, D.D. *Curr. Opin. Neurobiol.* **22**, 154–161 (2012).
7. Lowe, D.G. *Int. J. Comput. Vis.* **60**, 91–110 (2004).
8. Saalfeld, S., Cardona, A., Hartenstein, V. & Tomancak, P. *Bioinformatics* **26**, i57–i63 (2010).
9. Guest, E. & Baldock, R. *Bioimaging* **3**, 154–167 (1995).
10. Schmitt, O., Modersitzki, J., Heldmann, S., Wirtz, S. & Fischer, B. *Int. J. Comput. Vis.* **73**, 5–39 (2007).
11. Saalfeld, S., Cardona, A., Hartenstein, V. & Tomancak, P. *Bioinformatics* **25**, 1984–1986 (2009).
12. Cardona, A. *et al. PLoS ONE* doi:10.1371/journal.pone.0038011 (in the press).
13. Cardona, A. *et al. J. Neurosci.* **30**, 7538–7553 (2010).
14. Schindelin, J. *et al. Nat. Methods* (in the press).
15. Micheva, K.D. & Smith, S.J. *Neuron* **55**, 25–36 (2007).

ONLINE METHODS

The elastic model. We achieve globally minimized deformation by modeling alignment as a 2D elastic system of vertices connected by ideal springs according to Hooke's law. A Hookean spring has a relaxed length at which it exerts no force. Either extending or compressing a Hookean spring results in increasing stress. The stress amplitude is proportional to the difference of the spring's actual length and its relaxed length. Springs connecting the vertices of an 'image mesh' have a relaxed length corresponding to the distance between the vertices in the non-deformed image. Deforming the image mesh compresses and extends springs and therefore results in stress. Hooke's law enables us to model springs with a relaxed length of 0 for which no physical equivalent exists. A zero-length spring exerts force proportional to its extension beyond zero length; it cannot be compressed. Zero-length springs can be used to connect points that should be positioned at the same location. We connect corresponding locations between two overlapping images (tiles in a montage or sections in a section series) by zero-length springs. These springs aim to warp the images toward perfect overlap. In contrast, the nonzero-length springs within image meshes prefer a locally rigid transformation of each image. That way, the system penalizes arbitrary warp and distributes deformation evenly among all images.

Each image is tessellated into a mesh of regular triangles with each vertex connected to its neighboring vertices by a spring whose relaxed length is the original edge length of the triangle (**Fig. 1b** and **Supplementary Note**). For those vertices of the mesh on image I_1 overlapping image I_2, we identify their corresponding location in image I_2 by block matching. The vertex is then connected into the mesh on image I_2 by a zero-length spring with its target end located at an arbitrary place inside a triangle of the target mesh. Note that this 'passive' end does not contribute to the deformation of the mesh on image I_2 because it is not connected to any of its vertices by a spring. During relaxation, its location is updated according to the affine transformation defined by the three vertices of the embedding triangle. Vice versa, vertices of the mesh on image I_2 are connected to their corresponding location in image I_1, with their 'passive' ends updated according to the affine transformation of the embedding triangle in the mesh on image I_1.

The stiffness of ideal Hookean springs is specified by the spring constant k. Increasing the spring constant for springs spanning the triangle mesh will lead to less-deformed images and also less–well-aligned solutions. Using too-small spring constants effectively eliminates the elastic constraint and will therefore result in arbitrarily warped solutions. We have empirically estimated a spring constant $k = 0.1$ to be appropriate for our TEM series. During series alignment, the spring constant for cross-section springs depends on the index distance d in the series ($k = 1/d$), which gives farther sections less impact.

We relax the elastic system using an iterative solution similar to gradient descent. The desired end state of the system occurs when, for each vertex, the forces of all attached springs combine to equal 0. The force vector $\vec{F}$ for a vertex $\vec{p}_0$ can be calculated using Hooke's law (equation (1) and **Supplementary Fig. 5**).

$$\vec{F} = -\sum_{i\in\{1\ldots n\}} k_i \vec{x}_i \quad \text{with} \quad \vec{x}_i = \left(1 - \frac{l_i}{|\vec{p}_i - \vec{p}_0|}\right)(\vec{p}_i - \vec{p}_0) \quad (1)$$

At each iteration, force vectors are calculated for all vertices, and then all vertices are moved alongside their force vector. The distance of the move is the length of the force vector divided by the length of the largest force vector in the entire system. That way, the maximum step size per iteration is one pixel. All 'passive' spring ends are moved according to the affine transformation specified by the embedding triangle, which preserves their relative location in the triangle. The solution typically converges within a few hundred iterations.

Matching corresponding image content. Our method incorporates two techniques for establishing pairwise correspondences $(\vec{p}, \vec{q})$ between a point $\vec{p}$ in an image I_1 and a point $\vec{q}$ in an image I_2: (i) matching invariant local image features and (ii) matching blocks. Invariant local image features are used to establish sparse sets of corresponding landmarks between two images for which an approximate alignment is not known. We use the popular scale-invariant feature transform[7] for interest-point detection and feature matching. An approximate alignment for pairs or groups of overlapping images can be established by least-squares fitting an appropriately simplified transformation model (for example, a rigid transformation for each section; see **Fig. 1c**) to corresponding landmarks. In a previously published method[8], we estimated the optimal rigid transformation for each individual tile of a large tiled section series simultaneously; each tile connects to overlapping tiles within the section and across the series. Although it does not compensate for low-scale deformation and it delivers insufficiently stitched montages, the method can serve as a very good initialization for elastic montaging and series alignment of such data sets. We have extended this method to estimate an affine transformation per each tile that is regularized with respect to a rigid transformation, which effectively prevents arbitrary shear and scaling while better compensating for nonrigid deformation (**Supplementary Note**).

Block matching is performed on the approximately prealigned images. The local vicinity around each vertex of the section spring mesh is inspected for an optimal match. We use the normalized cross-correlation (NCC) coefficient r of a patch around the vertex and the overlapping patch in the other image as quality measure for a match. The location with maximal r specifies the offset of the vertex relative to the initial linear alignment. Block matching is executed on reasonably downscaled versions of the images. The ideal scaling factor depends on the application and quality of the signal. In our ssTEM series, the disparity between lateral and axial resolution suggests a scaling factor of 0.1 by which isotropic resolution is achieved. To overcome the reduced accuracy of the estimated offset, we use Brown's method[16] to calculate an approximate subpixel offset.

Filtering spurious matches. Both image-feature matching and block matching are local methods and can generate false positives. We reject those with a set of filters that exploit local (**Supplementary Fig. 3**) and global properties of the matches.

Correlation threshold: block matches with an NCC coefficient r below a user specified threshold are rejected (**Supplementary Fig. 3c**). The NCC coefficient ranges from −1.0 to 1.0 with $r = 1.0$ indicating perfect linear dependency, $r = 0.0$ indicating no linear dependency and $r = -1.0$ indicating inverse linear dependency.

Edge responses: block matches as well as interest points for feature detection may be detected on top of elongated structures (edges, ridges, stripes) and therefore poorly localized alongside the structure (**Supplementary Fig. 3d,f**). Such detections have a large (orthogonal to the ridge) and a small (alongside the ridge) principal curvature and can thus be identified by a large ratio between the two values[7]. Detections with a ratio larger than a given threshold are rejected.

Ambiguous matches: for feature descriptor matching, Lowe proposed comparing the distances of the reference to the second-best and best match[7]. For a distinctive true match, the ratio between the two distances is likely to be significantly lower than 1.0, whereas for a wrong match, many non-best distances are expected to be similar to the best match, thus leading to a ratio close to 1.0. During block matching, we use the filter to reject matches with multiple offsets, which results in a similar correlation (**Supplementary Fig. 3e,f**).

Geometric consensus: using local methods exclusively will lead either to false positives being accepted when using too-soft constraints or to many correct matches being rejected when using too-hard constraints. We therefore use the consensus of matches that were filtered by moderate local filters to reject the remaining outliers. The methods used are the random sample consensus[17] (RANSAC) and two variants of robust regression. All three methods make use of the observation that the hidden transformation is supported by all true matches up to an approximately normal distributed transfer error, whereas wrong matches do not support a common transformation. RANSAC identifies a hidden transformation by counting the supporting matches for many hypotheses generated from random minimal samples. If the minimal sample contains only true positives, then the hypothesis will be supported by all true positives. The best hypothesis is that which has the highest number of supporters. RANSAC is very effective for separating a small fraction of correct matches from a large set of false positives, but it leaves open the threshold for accepting a match as a supporter. That gap is closed by using a robust regression estimator that combines a least-squares estimator with an outlier filter based on error statistics in an iterative loop. It effectively removes moderate fractions of outliers while automatically estimating the required threshold.

Feature matches are filtered using RANSAC followed up by robust regression for a simple linear transformation model[8]. Block matches are filtered using another variation of robust regression. Each match $(\vec{p}_0, \vec{q}_0)$ is inspected individually to determine whether it is an outlier. To that end, all other block matches $(\vec{p}_i, \vec{q}_i)$ are used to estimate a linear transformation T (for example, an affine transformation) by means of weighted least squares with each match $(\vec{p}_i, \vec{q}_i)$ being weighted by a Gaussian radial distribution function (RDF) ω centered at the reference $(\vec{p}_0, \vec{q}_0)$ (equation (2)).

$$\arg\min_{T} \sum_{i \in \{1 \ldots n\}} \omega_i \left| T(\vec{p}_i) - \vec{q}_i \right|^2 \quad \text{with} \quad \omega_i = e^{\frac{\left| \vec{p}_i - \vec{p}_0 \right|^2}{-2\sigma^2}} \qquad (2)$$

Choosing a larger s.d. σ for the RDF ω requires the deformation field to be smoother. A match is rejected if its transfer error with respect to T is larger than a given threshold or if it is larger than

k times the average transfer error. The average transfer error is accumulated from all matches weighted by the RDF ω accordingly. The filter is applied in a loop until no match had been removed.

Naturally, the fraction of correct matches degrades with increasing distance of two sections in the series. It can therefore be used as a coarse deformation invariant distance metric to correct ordering mistakes and to estimate the approximate size of gaps in the series (**Supplementary Fig. 4**).

Manual skeleton traces. Jitter as introduced by insufficient alignment increases the total length of skeleton traces. We therefore report the scale-normalized total length l of the skeleton traces shown in **Figure 3** as an alignment quality criterion (equation (3) and **Supplementary Fig. 15**).

$$l = \sum_{\forall(\vec{p}, \vec{q})} \sqrt{\frac{(p_x - q_x)^2 + (p_y - q_y)^2}{s^2} + (p_z - q_z)^2} \qquad (3)$$

The total length l is the sum of all edge lengths. All edge lengths ($|\vec{p}, \vec{q}|$) are normalized by a local scale factor s that is the average scale factor of the contributing sections. The scale factor of a section is the average scale factor of all image tiles in the section, and the scale factor of an image tile is estimated through a least-squares approximation of its nonlinear elastic transformation by a similarity transformation (scale, rotation, translation). Shorter total length l implies improved alignment. Scale normalization makes the length measure invariant to global scaling. Without scale normalization, globally reducing the size of all (or a range of) sections would reduce the total length and render its applicability as a quality criterion for alignment useless. Using elastic alignment results in an l value 56.4% lower that that obtained with a rigid series alignment and 19.1% lower than with our previous method[8].

We compare the scale-normalized skeleton length l with a lower bound length f. The lower bound length f is the skeleton length after all edges between branch and end points have been replaced by straight lines (**Supplementary Fig. 16**). Because now only branch and end points suffer from alignment errors, the lower bound length f is robust with respect to insufficient section-to-section alignment. This robustness is reflected in the observation that elastic alignment decreases f by only 4.8% compared with a rigid series alignment and 0.7% compared with our previous method[8]. On the other hand, the percentage difference between l and f serves as an indicator for overall alignment quality. This difference is reduced to 31.8% by elastic alignment, compared with 188.0% by the rigid series alignment and 61.7% by our previous method[8], thus demonstrating the superior alignment results achieved by the elastic method. It is important to note that jitter in the manually generated skeleton traces comes not solely from insufficient alignment but also from inaccurate manual operation. This is particularly relevant for the annotations used in this paper because they were performed on poorly aligned data and annotation speed had a higher priority than accurate localization of each profile's center point. Our prediction is therefore that these skeleton traces cannot be used to report qualitative improvement over the current series alignment because the manual error already outweighs the alignment error.

Artificially generated ground truth. As suggested earlier[8], we have quantitatively evaluated the accuracy of our elastic alignment method using artificially generated ground truth. Using the open-source ray tracer POV-Ray (http://www.povray.org/), we have generated a synthetic volume that has the shape of a distorted ball filled with volumetric texture that resembles membranes and blob-like structures as present in biological tissue. We have artificially sectioned the volume at a section thickness of 2 pixels and generated two series of 400 sections, each 2,000 × 2,000 pixels. Evaluation series A repeats the same section 400 times (**Supplementary Video 2**). In this series, texture displacement is the exclusive result of deformation because no 'biological' signal changes occur alongside the z axis. Evaluation series B consists of 400 serial sections including the signal change induced by the volume (**Supplementary Video 3**) and as such is a more realistic test case. We have artificially distorted all sections of both series using randomized smooth nonlinear transformations using a moving least squares–affine transformation[18] for four control points at random source locations in either of the four quadrants of the image displaced by a maximum distance of 50 pixels. That induced section-to-section pairwise local deformation of up to 200 pixels relative to a rigid least-squares approximation. Each section was then rotated by a random angle and shifted in a random direction by up to 150 pixels. Both evaluation series were aligned using a rigid transformation per section, a regularized affine transformation per section (**Supplementary Note**) and our elastic alignment method on top of the affine alignment.

We report the average scale factor of each section relative to ground truth for all three alignment methods (**Supplementary Fig. 7**). Rigid series alignment per definition preserves the average section scale that has been introduced by nonlinear deformation. Both affine and elastic alignment recover the original scale of all sections across the entire series. The elastic method performs better as it can compensate for nonlinear deformation.

We compare alignment precision using a sample of straight lines projected along the z axis through the ground-truth series. Ideally, these lines should be reconstructed as straight lines along the z axis. Only points covered by the 'specimen' are considered because background is not expected to, and does not need to, be aligned. For all lines, we report the absolute displacement in the x, y plane relative to ground truth (**Supplementary Figs. 8**–**10**) and section-to-section pairwise displacement (**Supplementary Figs. 11** and **12**) in each z section. Ground truth and reconstruction results were previously aligned by a 2D rigid transformation to compensate for a global rotation and translational offset. Elastic alignment clearly outperforms rigid and affine alignment in its ability to recover the original shape of the 'specimen' while at the same time effectively removing section-to-section jitter.

16. Brown, M. & Lowe, D. in *British Machine Vision Conf.* 656–665 (BMVC, (2002)).
17. Fischler, M.A. & Bolles, R.C. *Commun. ACM* **24**, 381–395 (1981).
18. Schaefer, S., McPhail, T. & Warren, J. *ACM Trans. Graph.* **25**, 533–540 (2006).