Bead-based Mosaicing of Single Plane Illumination Microscopy Images Using Geometric Local Descriptor Matching

Stephan Preibisch[a], Stephan Saalfeld[a], Torsten Rohlfing[b] and Pavel Tomancak[a]

[a]Max Planck Institute of Molecular Cell Biology and Genetics, 01307 Dresden, Germany
[b]SRI International, Neuroscience Program, Menlo Park, CA 94025, USA

ABSTRACT

Single Plane Illumination Microscopy (SPIM) is an emerging microscopic technique that enables live imaging of large biological specimens in their entirety. By imaging the biological sample from multiple angles, SPIM has the potential to achieve isotropic resolution throughout relatively large biological specimens. For every angle, however, only a shallow section of the specimen is imaged with high resolution, whereas deeper regions appear increasingly blurred. Existing intensity-based registration techniques still struggle to robustly and accurately align images that are characterized by limited overlap and/or heavy blurring. To be able to register such images, we add sub-resolution fluorescent beads to the rigid agarose medium in which the imaged specimen is embedded. For each segmented bead, we store the relative location of its n nearest neighbors in image space as rotation-invariant geometric local descriptors. Corresponding beads between overlapping images are identified by matching these descriptors. The bead correspondences are used to simultaneously estimate the globally optimal transformation for each individual image. The final output image is created by combining all images in an angle-independent output space, using volume injection and local content-based weighting of contributing images. We demonstrate the performance of our approach on data acquired from living embryos of *Drosophila* and fixed adult *C.elegans* worms. Bead-based registration outperformed intensity-based registration in terms of computation speed by two orders of magnitude while producing bead registration errors below 1 µm (about 1 pixel). It, therefore, provides an ideal tool for processing of long term time-lapse recordings of embryonic development consisting of hundreds of time points.

Keywords: Registration, Image Restoration and Enhancement, Single Plane Illumination Microscopy, Local Descriptors

1. INTRODUCTION

1.1 Biological Background

Early animal development can be described as a complex, three-dimensional cellular system that changes dramatically across time as a consequence of cell proliferation, growth, differentiation, and movements. In order to understand these developmental processes it is essential to image developing biological samples, such as embryos, with sufficient spatial resolution to distinguish individual cells throughout the entire sample volume and with sufficient temporal resolution to follow morphogenetic events during development. Imaging of specimen in their entirety (*in toto*) cannot be achieved by current state-of-the-art microscopic techniques such as confocal, 2-photon or spinning disc microscopes, as they lack spatial resolution throughout the specimen and can achieve high temporal resolution only in very small volumes.

Single Plane Illumination Microscopy (SPIM)[1] is an emerging technique that enables *in toto* live imaging of large specimens with high temporal and spatial resolution. By imaging the sample from multiple angles, herein referred to as *views*, SPIM has the potential to achieve isotropic resolution throughout relatively large specimens. This property is not only interesting for live imaging but also for very large* fixed specimen such as adult *Caenorhabditis elegans (C. elegans)* worms. The necessary prerequisite for realizing the full potential of SPIM for biological imaging, however, is an image mosaicing step that achieves precise registration of the data acquired from multiple angles, optionally followed by fusion into a single, isotropic volume.

We use SPIM recordings of *Drosophila* embryogenesis and fixed *C. elegans* samples to test the performance of the algorithms we have developed for this purpose. The *Drosophila* embryo is a premium model system for studying animal embryonic processes. An extensive reverse genetic toolkit allows us to mark general morphological features such as cell nuclei in the embryo as well as to generate specific gene expression reporters for any gene in the fruit fly genome.

---

*The term large always refers the size of the imaged specimen relative to the magnification used.
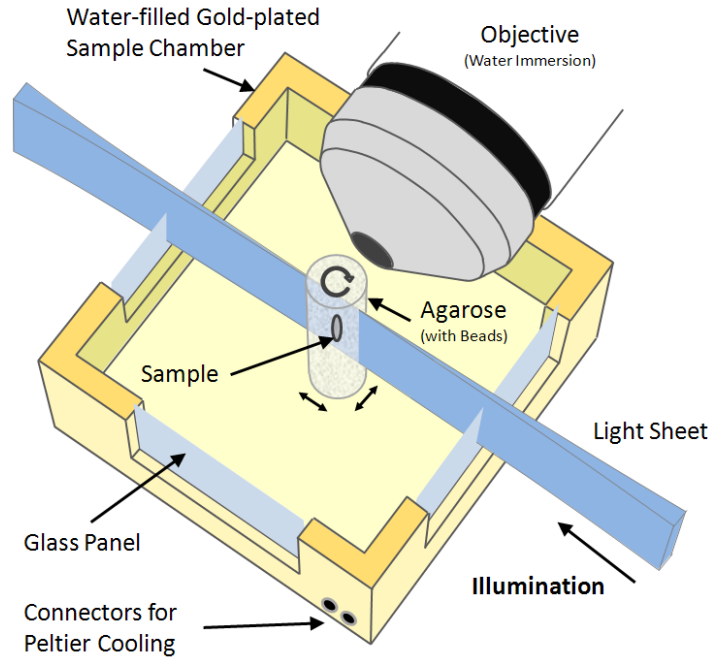
Figure 1. Schematic drawing of SPIM acquisition setup. The specimen is embedded in agarose gel protruding from a glass capillary (not shown). The excitation laser (blue) is formed into a light sheet and the emitted fluorescent light is detected by a CCD camera through an exchangeable water dipping objective arranged perpendicular to the axis of illumination. 3D stacks are acquired by moving the specimen through the light sheet. The capillary is rotated and the light sheet illuminates the specimen at different angles. The drawing is adapted from the PhD thesis of Dr. Jan Huisken.[2]

Using SPIM, we intend to map the dynamic spatial and temporal changes in gene expression specificity for all important developmental genes with cellular resolution in the context of the entire, intact developing embryo. Our algorithms for SPIM mosaicing are crucial for efficient processing of these systematically acquired *Drosophila* embryo recordings, but are also applicable for processing of any multi-view SPIM image data of other biological specimens or even very small medical samples.

2. METHODS

2.1 Imaging

A Single Plane Illumination Microscope, illustrated in Figure 1, achieves optical sectioning by focusing the excitation laser into a thin, concave laser light sheet that reaches its minimal thickness in the center of detection. Depending on the magnification, the thickness of the light sheet varies; higher magnifications require thinner light sheets, which also exhibit significantly increased concavity. It enters the water-filler sample chamber through a thin glass panel and illuminates the sample, which is embedded in agarose gel protruding from a glass capillary. The capillary is attached to a motor that rotates the sample. The objective lens is mounted perpendicular to the axis of illumination so that the laser only illuminates the plane of the specimen that the lens is focused on, thus minimizing photo-bleaching and laser damage of the living samples. Two-dimensional images of emitted fluorescent light are captured by a CCD camera and discretized with a dynamic range of 8 or 12 bits. The use of a CCD camera enables a very fast acquisition rate, which is important for capturing dynamic developmental events in living embryos. In order to acquire 3D image stacks, the sample is moved perpendicular to the light sheet, in our case in increments of $1\,\mu m$ to $2\,\mu m$ depending on the objective used and the corresponding light sheet thickness.

The SPIM instrument can, in principle, achieve an isotropic, high resolution along x, y and z axis allowing for the first time *in toto* imaging of large three-dimensional specimens. In order to achieve isotropic resolution uniformly across the sample volume, it is necessary to rotate the sample and record image stacks for the same specimen from different angles (usually 4 to 12).

Herein, we present results from imaging two different specimen. Firstly, we imaged a developing *Drosophila* embryo expressing fluorescent HIS-YFP[†] marker in all nuclei using a prototype SPIM microscope equipped with a Zeiss 20×/0.5 Achroplan objective. Secondly, we imaged a fixed *C. elegans* adult worm stained with Cytox-Green[‡] using a 40×/0.5 Achroplan objective. We incorporated small amounts of yellow fluorescent beads with a diameter of $0.5\,\mu m$ into the agarose medium before mounting to reach final concentration of about 1000-5000 beads per imaged volume, depending on the thickness of the sample. We acquired views at 8 angles spaced by 45° for each time point in the extended time-lapse recording, and 10 angles spaced by 36° for the the fixed sample.

2.2 Previous work & Motivation

Our previous work on SPIM data reconstruction investigated intensity-based registration methods. We applied groupwise congealing-based image registration[3] as well as pairwise Fourier transform-based image alignment.[4] Both used similar application-specific groupwise transformation models that incorporated per-image as well as groupwise transformation parameters, which significantly reduced the degrees of freedom by limiting the transformations to physically possible constellations.

In short, the "congealing" method puts all views into an empty reference coordinate space and calculates the entropy of pixel intensities over all views at each pixel, which is then summed over all pixels. By minimizing the total entropy, simultaneous registration of the entire group of images can be achieved. The Fourier approach determines the transformation parameters by maximizing the alignment quality of a translation-only registration between all the views for different hypotheses of the orientation of the common rotation axis. Unlike gradient-based methods the phase correlation[5] approach cannot get trapped in local minima, as it is a global approach that examines all possible translations in one operation. Although both intensity-based method proved useful they still suffer from three problems:

1. When imaging living specimen the sample is developing during acquisition of the views belonging to one time point. Acquisition takes roughly 30 seconds per view for a *Drosophila* embryo at 20×, which adds up to a few minutes per time point. During stages of fast developmental changes, such as gastrulation, an intensity-based registration becomes practically impossible as the specimen changes too dramatically before all views can be acquired.

2. When imaging with only a few angles at one time point to increase the temporal resolution, the overlap between neighboring views is reduced to a minimal extent rendering intensity-based registration problematic.

3. When processing time-lapse data with hundreds of time points the computation time becomes impractically high.

Those drawbacks encouraged us to develop a feature-based registration scheme that depends on sub-resolution fluorescent beads added to the agarose. They function as universal automatically-detectable landmarks and facilitate a sample-independent registration in a fraction of time necessary for intensity-based registration.

2.3 Image Registration

The registration of the views relies on matching of the fluorescent beads and consists of three steps. We first segment the beads in each view, then identify corresponding beads between all views and finally use all identified correspondences to find the globally optimal transformation for each view.

2.3.1 Bead Segmentation

Segmentation of the beads is the first step in the registration pipeline. We have developed two methods for detection of the sub-resolution beads: a simple threshold segmentation and a scale space-based bead identification.

[†]HIS = Histone, acts as spool for DNA in the nuclei ; YFP = Yellow Fluorescent Protein
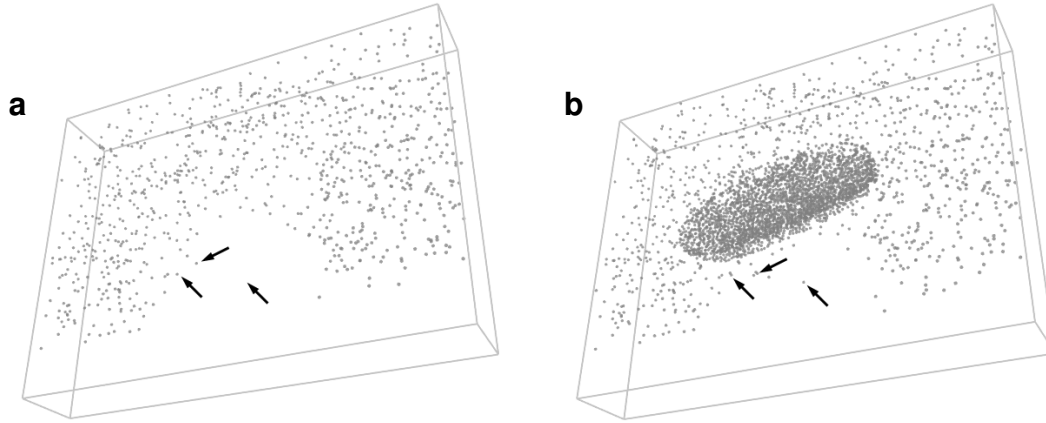[‡]Cytox-Green is a green fluorescent dye for fixed samples that stains DNA

Figure 2. Shown here is a 3D visualization of the segmented beads in one view of an acquisition of a *Drosophila* embryo, see Fig. 9 for a 3D visualization of the input image. (a) shows the result of the threshold segmentation, in which 901 beads were found, (b) shows the result of the scale space-based segmentation, in which 3405 bead-like objects were identified. The arrows mark exemplary beads which were missed by the threshold segmentation. The scale space-based segmentation oversegments the image and detects most of the nuclei of the imaged *Drosophila* sample. In both segmentation results one can clearly see the "shadow" cast by the embryo as well as the edge of the agarose that the specimen is embedded in. All 3D visualizations shown in this publication were created using the 3D viewer of Benjamin Schmidt.

Threshold Segmentation. The beads are segmented using a very low threshold (2% of the maximal intensity) on the raw SPIM image that has very little noise in the background. The beads have a diameter of approximately $0.5\,\mu m$, and the pixel size of the acquired images is $0.73 \times 0.73\,\mu m$ and $0.365 \times 0.365\,\mu m$ for the 20× and the 40× objective, respectively. Therefore the beads are expected to appear as the point spread function of the microscope in the image. We apply size and shape constraints to each connected component to remove touching beads and the segmentations within the volume occupied by the imaged sample. The resulting image of each bead in 3D space depends both on the point spread function and the number of z planes in which it is hit by the light sheet; a property which does not have to be symmetric in the axial direction. Therefore the maximum intensity pixel in each segmented area is used as the estimate of the bead position rather than the center of mass.

Scale Space Segmentation. The scale space-based approach relies on the assumption that the beads can be detected at low scales in the Difference-of-Gaussian (DoG) space.[6,7] To extract the beads we compute the DoG images $D(x, y, z, \sigma)$ of the lowest octave and extract all local minima in a 3×3×3×3 neighborhood (x, y, z, σ) as they represent intensity maxima of the respective size σ in the image space. The sub-pixel location of each detection is then computed by fitting a four-dimensional quadratic function to the sample points, which is a straightforward extension of the three-dimensional case presented by Brown and Lowe.[8] In order to find all beads we have to detect even the smallest features in the image. Lowe[7] suggests to initially increase the image size by a factor of two, and thus add another octave for very small scales. Because of the memory and time consumption this is impractical for 3D images with sizes of 1388×1040×100–400 pixels. Instead, we analyze the minima of the lowest step of the octave in a 3×3×3×**2** neighborhood, which detects the beads reliably.

The threshold-based segmentation finds most of the beads. The computation time increases linearly with increasing axial size of the image but it takes only seconds to compute even on a typical notebook. However, it misses beads which are close to each other or to the sample, or those where the segmented area has an unexpected shape. On the other hand, the scale-space approach massively oversegments the image to find all beads (verified manually, data not shown). Additionally many "blob-like" structures inside the imaged sample are detected, but these do not disturb the registration process. Memory and computation time requirements are significantly increased, relative to threshold segmentation, and depend cubically on the image size. Figure 2 visualizes the results of both segmentation strategies. All further data shown in this publication are based on the scale space-based segmentation.
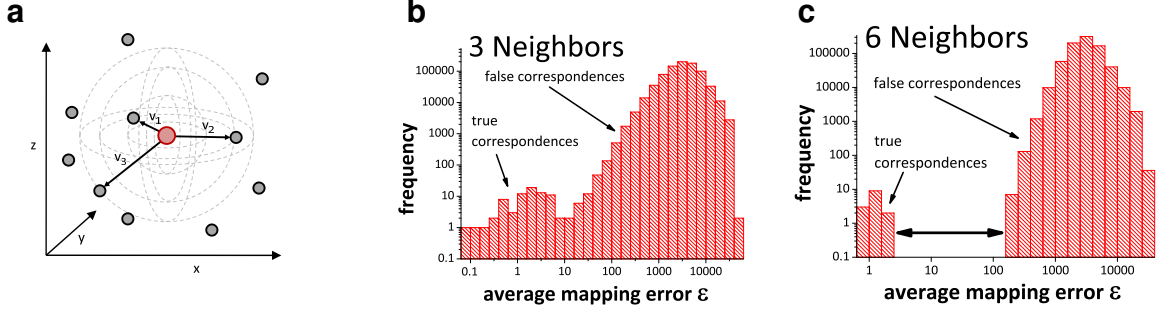
Figure 3. Principle of the geometric point descriptor (a). We detected 917 and 881 beads in two neighboring views. After the all-to-all matching using 3 neighbors, true and false correspondences are not easily separable (b), whereas using 6 neighbors true correspondences are clearly separated from the false correspondences (c). With RANSAC[13] we identified 71 true correspondences using 3 neighbors and 14 using 6 neighbors.

2.3.2 Establishing Bead Correspondences

Corresponding bead pairs $(\vec{a}, \vec{b})$ between two views A and B have to be identified invariantly to translation and rotation. To this end, we have developed a *geometric local descriptor* inspired by work done in object recognition.[9, 10] The local descriptor of a bead is defined by the locations of its n nearest neighbors in 3D image space, ordered by their distance to the bead (see Fig. 3a). Translation invariance is guaranteed by storing locations relative to the bead. That is, each bead descriptor is an ordered 3D point cloud of cardinality n with its origin $\vec{o} = (0, 0, 0)^{\mathsf{T}}$ being the location of the bead. Nearest neighbors in image space are identified using the kd-tree implementation of the WEKA framework.[11]

Local descriptor matching is performed invariantly to rotation by mapping the ordered point cloud of bead $\vec{a} \in A$ to that of bead $\vec{b} \in B$ by means of least squares-error point mapping using the closed-form unit quaternion-based solution.[12] The similarity measure ϵ is the average point mapping error. Each candidate in A is matched against each candidate in B. Corresponding beads are those with minimal ϵ. Typically, many beads visible in view A are not visible in view B. Correspondence matching as described above will identify false correspondence candidates for these beads. We found empirically that true correspondence candidates have a significantly lower ϵ than false positives (see Fig. 3c). To reject false positives, we first apply an adaptive threshold on ϵ that ensures the best match for a geometric local descriptor to be at least one order of magnitude better than the second-best match.[7] We found that the distinctiveness of the local descriptor increases with the number of neighbors used, at the expense of fewer identified correspondences (see Fig. 3b, c). To overcome this problem, and to identify the maximum number of true correspondences, we reject false correspondences from candidate sets of the minimal descriptor size ($n = 3$ neighbors for rigid) with the Random Sample Consensus (RANSAC)[13] on a rigid or affine transformation model. To ensure the validity of the identified set of correspondences between view A and view B it must contain at least 3 times as many correspondences as is the minimal number required for the used transformation model.

2.3.3 Groupwise registration

The identified set of corresponding beads $C_{AB} = \{(\vec{a}_i, \vec{b}_i) : i = \{1, 2, \ldots, |C|\}\}$ for a pair of views A and B defines a rigid or affine transformation $\mathbf{T}_{AB}$ that maps A to B by means of least square bead correspondence displacement

$$\arg\min_{\mathbf{T}_{AB}} \sum_{(\vec{a}, \vec{b}) \in C_{ab}} \left\| \mathbf{T}_{AB}\vec{a} - \vec{b} \right\|^2.$$

For this problem, we use an affine transformation or employ the closed-form unit quaternion-based solution[12] as used for local descriptor matching.

More than two views require groupwise optimization of the configuration $T_{VF} = \{\mathbf{T}_{AF} : A, F \in V\}$ with V being the set of all views and F being a fixed view that defines the common reference frame. Then the above equation extends to

$$\arg\min_{T_{VF}} \sum_{A \in V \setminus \{F\}} \left(\sum_{B \in V \setminus \{A, F\}} \left(\sum_{(\vec{a}, \vec{b}) \in C_{AB}} \left\| \mathbf{T}_{AF}\vec{a} - \mathbf{T}_{BF}\vec{b} \right\|^2 \right) \right)$$
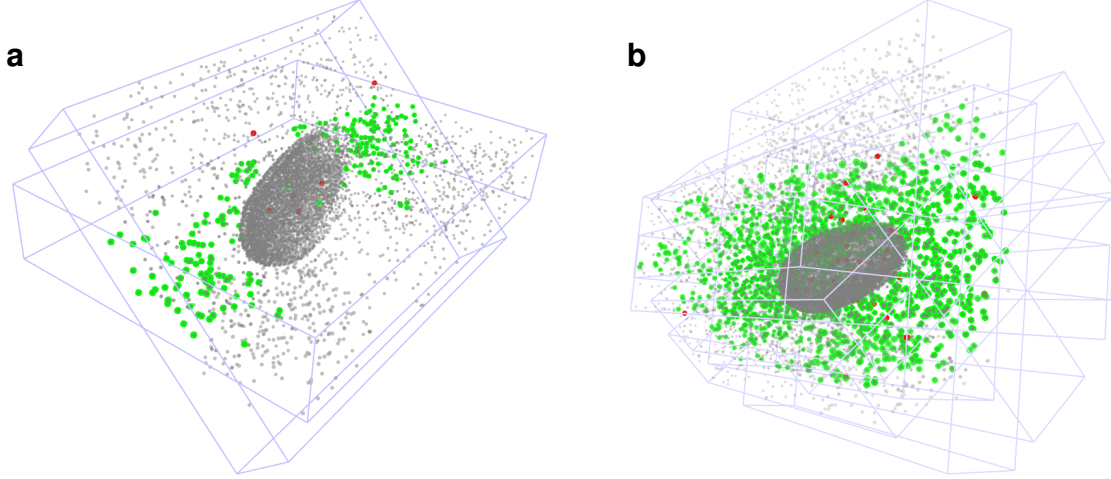
Figure 4. Shown here is a 3D visualization of the bead registration. All colored beads (green and red) are detected as correspondences by the rigid local descriptor matching. The green colored beads are true correspondences, identified as those who match the same affine transformation between the views computed by the RANSAC. The red beads were identified by RANSAC to be wrong correspondences. The relative proportion of wrong correspondences is typically in the range of 0–2%. (a) visualizes the registration for two adjacent views, between which 237 true and 7 false correspondences were detected using an affine model. (b) shows the registration for all views at once, where 1100 true and 37 false correspondences were detected using an affine model.

with C_{AB} being the set of bead correspondences $(\vec{a}, \vec{b})$ between view A and view B with $\vec{a} \in A$ and $\vec{b} \in B$. This problem is solved using an iterative optimization scheme. In each iteration, the optimal transformation $\mathbf{T}_{AF}$ for each single view $A \in V \setminus \{F\}$ relative to the current configuration is estimated and applied to all beads in this view. The scheme terminates on convergence of the overall bead correspondence displacement.
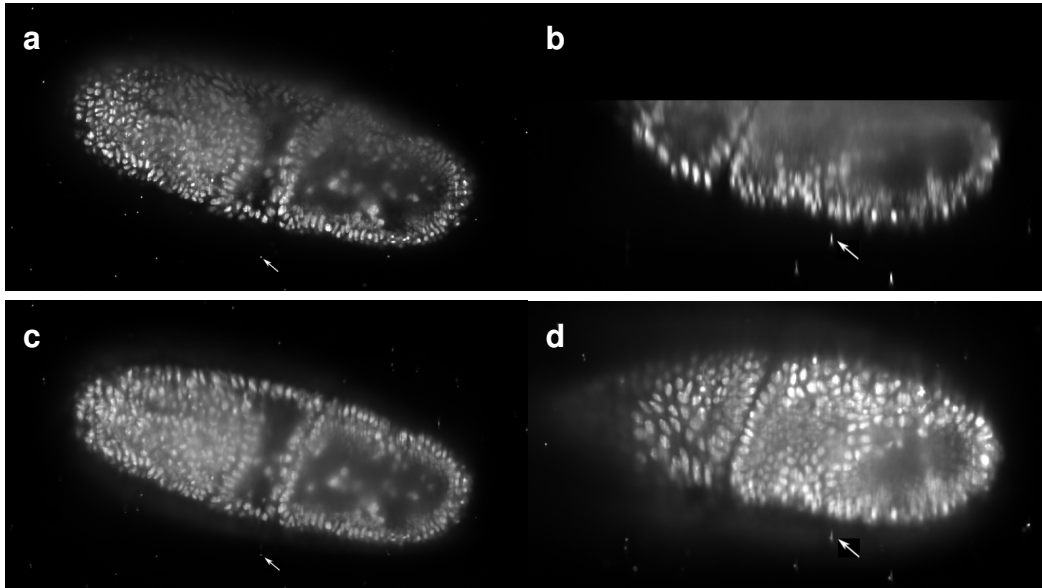

Figure 5. Result of bead-based SPIM data mosaicing. (a) xy-plane and (b) xz-plane through the same single SPIM view of the *Drosophila* embryo and (c,d) through the volume reconstructed from 8 views (resampled to $0.73 \times 0.73 \times 0.73\,\mu$m isotropic pixel size). The mosaicing restores the blurry upper part of the single view data (a,c), increases the axial resolution (b,d) and restores from other views the parts that were left out during acquistion (b,d). The arrows mark two exemplary fluorescent beads.
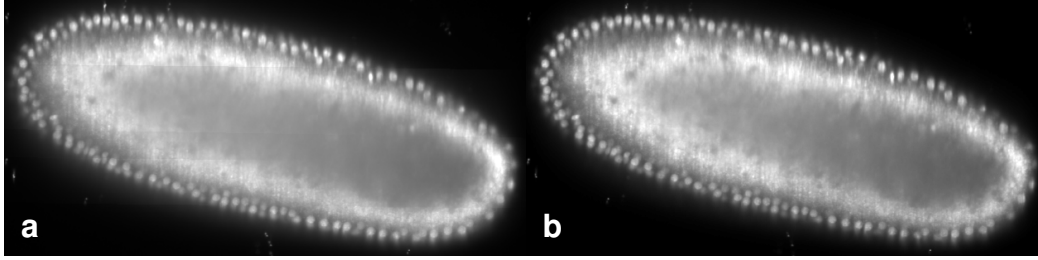
Figure 6. Results of the non-linear blending during image fusion. (a) shows one slice through the reconstructed sample and the line artifacts (horizontal lines) introduced by incomplete scanning of the specimen in axial direction in each view. (b) shows the same slice, but with additional blending applied. The boundaries of the individual views are not noticeable anymore.

2.4 Image Fusion

An effective fusion algorithm of the aligned image stacks must ensure that each view contributes to the final fused volume only useful sharp image data acquired from the area of the sample close to the detection lens. We use Gaussian Filters as an approximation of the local image information and use it as a weightening factor for each pixel in each contributing views.[3] We fuse the final output image using volume injection[14] to account for resolution differences in the axial and lateral dimensions.

For each view we typically do not image the entire specimen depth, but stop at roughly two thirds of its thickness, as the acquired data becomes increasingly blurred (see Fig. 5b). However, this introduces line artifacts when reconstructing the final output image (see Fig. 6a). We suppress these using nonlinear blending[15] between the views, an approach similar to the linear blending proposed by Brown and Lowe.[16] We compute the blending in overlapping areas by computing a weighting factor for each contributing view at each location, which is based on the distance of the pixels from the view boundary. Figure 6 illustrates the result of the blending process.

3. RESULTS

Post-processing of SPIM data requires highly accurate registration in order to enable accurate segmentation and tracing of biologically interesting entities such as cells and tissues. We demonstrate the accuracy of the bead-based mosaicing approach by reconstructing acquisitions of *Drosophila* and *C.elegans*. The eight-view acquisition of a *Drosophila* embryo imaged with a 20× objective ($0.73 \times 0.73 \times 0.73$ μm isotropic pixel size, see Fig. 5) results in a final output image with $1416 \times 1230 \times 1076$ px and produces an average bead registration error of 1.8 px (min. 1.5 px, max. 2.0 px) when using
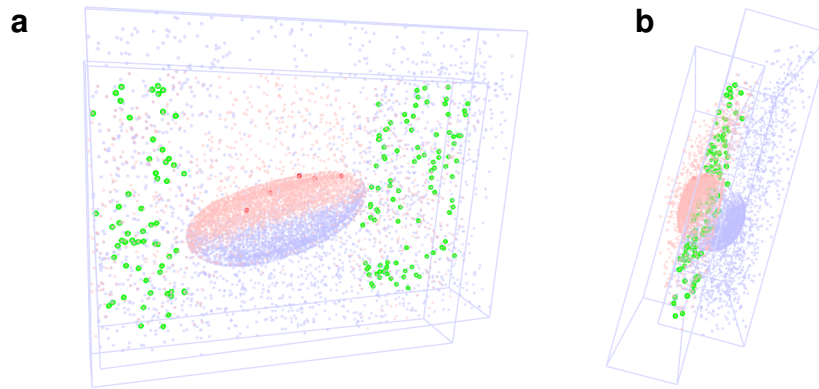

Figure 7. Shown here is a 3D visualization of the bead registration of two opposing views, 0° and 180°. The beads from 0° are colored in red, from 180° in blue. (a) shows how the shadow cast by the imaged specimen prevents the detection of corresponding beads in the center of the views: in the upper part of the image there are mostly red beads, in the lower part blue ones. (b) illustrates how corresponding beads are only found in the shallow overlapping area. 161 true and 5 false correspondences were identified although there is almost no overlap in image information as shown by the incidental segmentation of the nuclei.
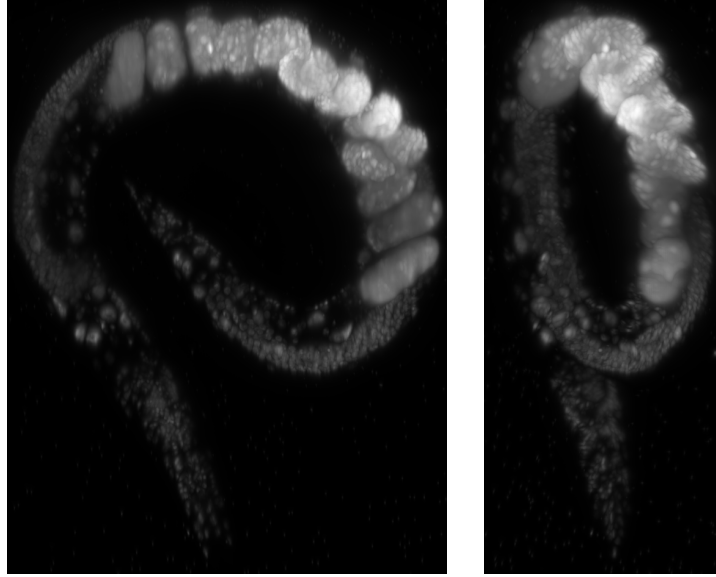
Figure 8. Shown here are two maximum intensity projections of a reconstructed 12-angle acquisition of *C.elegans* at 40× magnification. The images were fused into an isotropic 0.365×0.365×0.365 µm/px resolution output image. Note that the beads are almost invisible in this image, as we used yellow fluorescent beads in combination with a green sample staining.

a rigid model and an average error of 1.0 px (min. 0.9 px, max. 1.1 px) when using an affine model. When reconstructing the 40× ($0.365 \times 0.365 \times 0.365$ µm isotropic pixel size) 12-angle acquisition of the *C.elegans* adult worm, the rigid model is not able to produce a registration with an error lower than 5 px. The affine model produces an average bead registration error of 1.17 px (min. 1.05 px, max. 1.26 px) (see Fig. 8). The affine model is apparently able to match the beads with higher accuracy and precision than the rigid model. We speculate that this is due to the ability of the affine transformation to account for imaging effects which violate the rigidity assumption, such as minor misalignments of the SPIM system.

Figure 9 shows volume renderings of a reconstructed time series imaged with a 20× objective consisting of 42 time points. On average it produced an registration error of 1.82 ± 0.12 px (min. 1.55 ± 0.11 px, max. 2.08 ± 0.16 px) using a rigid model showing the stability of the bead-based approach.

To capture dynamic developmental processes with SPIM it is necessary to acquire multi-view data for hundreds of time-points generating massive datasets. The mosaicing methods, therefore, must be not only precise but fast as well. Computation time for a single time point of 20× *Drosophila* acquisition takes 10 minutes on an Intel® 4 CPU machine with 2.67 GHz and 24 GB of RAM. In contrast, alternative intensity-based methods using congealing[3] and iterative rotation axis estimation[4] takes about 4 and 8 hours, respectively. The bead-based method outperforms even the time-lapse optimized variant of the intensity-based method that uses prior rotation axis estimates for faster processing[4] by an order of magnitude. The computation of the time series is typically done on a cluster with relatively slow nodes and not more than 2GB of RAM. The computation time for the whole 42 time points on such a cluster was roughly 6 hours, which renders this approach suitable for long-term time lapse processing.

Furthermore, in order to maximize acquisition frame rate, SPIM views are often limited to the surface portion of the specimen, where high-content information can be recorded, and exclude the increasingly blurred regions deeper inside the specimen. The bead-based method performs optimally on such acquisitions (see Fig. 5b, d) while intensity-based methods often fail due to insufficient overlap of the views (see Fig. 7). Moreover, blending applied during the fusion stage eliminates artifacts arising from sharp edge differences in image intensity between incomplete views (see Fig. 5b, d).

4. DISCUSSION

To the best of our knowledge, this paper makes the following contributions: (1) We introduce the *geometric local descriptor* for reliable and fast matching of partially overlapping point clouds. (2) We describe an integrated algorithm for SPIM
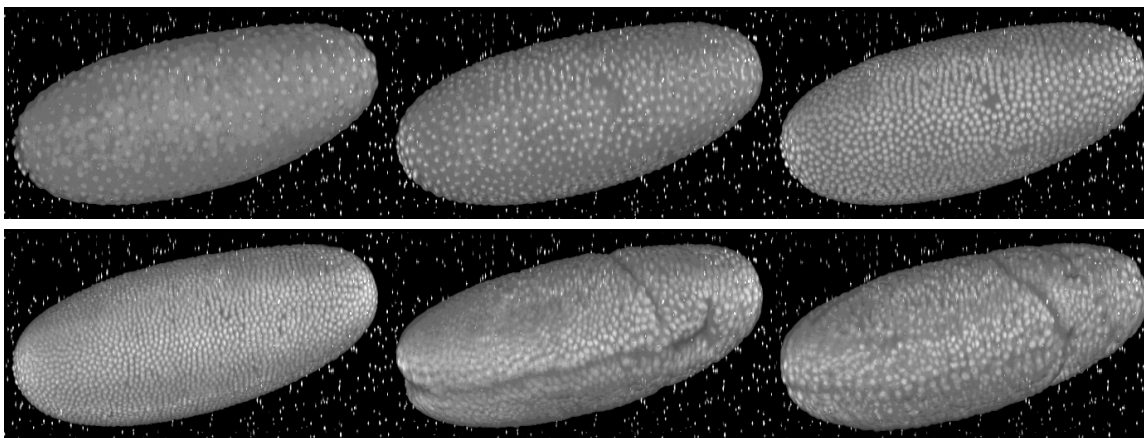
Figure 9. 3D Visualization of *Drosophila* time lapse registration. The images show six reconstructed time points covering early *Drosophila* embryonic development from blastoderm stage to gastrulation. The last two images appear more blurred as this marks the fastest stage of development, and the cells move significantly during acquisition of the images.

mosaicing that is able to register multi-view SPIM data orders of magnitude faster then existing methods and produces isotropic output images free of artifacts even when some of the views are incomplete.

In summary, we have introduced an image mosaicing technique for SPIM images that uses bead-based geometrical local descriptor matching and content-based image fusion. It is significantly faster than intensity-based registration algorithms[3,4] and is able to register datasets where intensity-based registration fails. The geometric local descriptor is able to identify even the smallest subsets of corresponding beads in partially overlapping views, and the groupwise registration produces a globally consistent arrangement of the views based on corresponding bead positions. Therefore the registration is completely independent of the biological image content, so that even very fast-changing samples or very uniform specimens can be registered, which makes this method a very generic and powerful tool for SPIM image acquisition. The bead based registration offers another advantage over the intensity-based registration, namely that there is a detection criterion for registration success based on a sufficient number of true bead correspondences for each view. When processing time lapses, this is a very useful property, as it eliminates the need for manual inspection of every single time point. Apart from the data shown in this paper we applied the method successfully to various other samples including *Zebrafish*, *Mus Musculus*, and human placenta samples.

The output image is created by combining volume injection, estimation of local image content, and nonlinear blending. The resulting images provide an ideal basis for subsequent non-rigid alignment, or they can be used directly for segmentation of nuclei in static specimens. The speed of the algorithms allows for the first time fast reconstruction of long-term time-lapse high-resolution SPIM data and even reconstruction of SPIM acquisitions on notebooks and normal workstation PCs.

Future work on the registration technique comprises generic non-rigid refinement of the registration to be able to better cope with small sample changes during time point acquisition and optical artifacts (scattering, diffraction) introduced by the microscope. For image fusion we want to include multi-band blending,[17] light simulations[18] to generate quality maps that are based on the imaging process, and deconvolution to increase the quality of the output images.

REFERENCES

1. J. Huisken, J. Swoger, F. D. Bene, J. Wittbrodt, and E. H. K. Stelzer, "Optical sectioning deep inside live embryos by selective plane illumination microscopy," *Science* **305 (5686)**, pp. 1007–1010, 2004.

2. J. Huisken, "Multiview microscopy and multibeam manipulation for highresolution optical imaging," phd thesis, Albert-Ludwigs-Universität Freiburg, 2004.

3. S. Preibisch, T. Rohlfing, M. P. Hasak, and P. Tomancak, "Mosaicing of single plane illumination miscroscopy images using groupwise registration and fast content-based image fusion," in *Medical Imaging 2008: Image Processing*, J. M. Reinhardt and J. P. W. Pluim, eds., *Proceedings of SPIE* **6914**(1), p. 69140E, 2008.

4. S. Preibisch, R. Ejsmont, T. Rohlfing, and P. Tomancak, "Towards digital representation of drosophila embryogenesis," in *Proceedings of 5th IEEE International Symposium on Biomedical Imaging*, pp. 324–327, 2008.

5. C. D. Kuglin and D. C. Hines, "The phase correlation image alignment method," in *International Conference on Cybernetics and Society*, *Proc. IEEE*, pp. 163–165, 1975.

6. T. Lindeberg, "Scale-space theory: A basic tool for analysing structures at different scales," *Journal of Applied Statistics* **21(2)**, pp. 224–270, 1994.

7. D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *IJCV* **60(2)**, pp. 91–110, 2004.

8. M. Brown and D. Lowe, "Invariant features from interest point groups," in *In British Machine Vision Conference*, pp. 656–665, 2002.

9. A. Frome, D. Huber, R. Kolluri, T. Bulow, and J. Malik, "Recognizing objects in range data using regional point descriptors," in *Proceedings of the European Conference on Computer Vision (ECCV)*, May 2004.

10. D. Bradley, D. Silver, and S. Thayer, "A regional point descriptor for global localization in subterranean environments," in *IEEE conference on Robotics Automation and Mechatronics (RAM 2005)*, **1**, pp. 440–445, December 2004.

11. I. H. Witten and E. Frank, *Data Mining: Practical machine learning tools and techniques*, Morgen Kaufmann, San Francisco, second ed., 2005.

12. B. K. P. Horn, "Closed-form solution of absolute orientation using unit quaternions," *Journal of the Optical Society of America A* **4(4)**, pp. 629–642, 1987.

13. M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM* **24(6)**, pp. 381–395, 1981.

14. F. Rousseau, O. Glenn, O. B. Iordanova, C. Rodriguez-Carranza, D. Vigneron, J. Barkovich, and C. Studholme, "A novel approach to high resolution fetal brain mr imaging," in *MICCAI*, **3749**, pp. 548–555, 2005.

15. S. Preibisch, S. Saalfeld, and P. Tomancak, "Fast stitching of large 3d biological datasets," in *Proceedings of the ImageJ User and Developer Conference*, **to appear**, 2008. preprint `http://fly.mpi-cbg.de/~preibisch/pubs/imagejpaper2008.pdf`.

16. M. Brown and D. G. Lowe, "Automatic panoramic image stitching using invariant features," *International Journal of Computer Vision* **74**, pp. 59–73, August 2007.

17. P. J. Burt and E. H. Adelson, "A multiresolution spline with application to image mosaics," *ACM Transactions on Graphics* **2**, pp. 217–236, 1983.

18. W. Wein, M. Blume, U. Leischner, H. U. Dodt, and N. Navab, "Quality-based registration and reconstruction of optical tomography volumes," in *MICCAI*, **4792**, pp. 718–725, 2007.