

Multi-view image fusion improves resolution in three-dimensional microscopy

Jim Swoger^{1*}, Peter Verveer², Klaus Greger, Jan Huisken³, and Ernst H.K. Stelzer

Cell Biology & Biophysics Unit, European Molecular Biology Laboratory (EMBL), Meyerhofstrasse 1, D-69117 Heidelberg, Germany.

¹Current affiliation: Systems Biology Programme, Centre for Genomic Regulation, C/ Dr. Aiguader 88, 08003 Barcelona, Spain.

²Current affiliation: Department of Systemic Cell Biology, Max Plank Institute of Molecular Physiology, Otto-Hahn-Strasse 11, 44227, Dortmund, Germany.

³Current affiliation: Dept. of Biochemistry & Biophysics, University of California, San Francisco, 1550 Fourth St., San Francisco, CA 94143-2711, U.S.A.

*Corresponding author: jim.swoger@crg.es

Abstract: A non-blind, shift-invariant image processing technique that fuses multi-view three-dimensional image data sets into a single, high quality three-dimensional image is presented. It is effective for 1) improving the resolution and isotropy in images of transparent specimens, and 2) improving the uniformity of the image quality of partially opaque samples. This is demonstrated with fluorescent samples such as *Drosophila melanogaster* and Medaka embryos and pollen grains imaged by Selective Plane Illumination Microscopy (SPIM). The application of the algorithm to SPIM data yields high-resolution images of organ structure and gene expression, in some cases at a sub-cellular level, throughout specimens ranging from several microns up to a millimeter in size.

©2007 Optical Society of America

OCIS codes: (180.0180) Microscopy; (180.6900) Three-dimensional microscopy; (100.0100) Image Processing; (100.3010) Image reconstruction techniques; (100.6890) Three-dimensional image processing.

References and links

1. Hell, S. & Stelzer, E.H.K. "Properties of a 4Pi confocal fluorescence microscope" *J. Opt. Soc. Am. A* **9**, 2159-2166 (1992).
2. Stelzer, E.H.K. & Lindek, S. "Fundamental reduction of the observation volume in far-field light microscopy by detection orthogonal to the illumination axis: confocal theta microscopy" *Opt. Commun.* **111**, 536-547 (1994).
3. Drexler, W. "Ultrahigh-resolution optical coherence tomography" *J. Biomed. Opt.* **9**, 47-74 (2004).
4. Gustafsson, M.G.L., Agard, D.A. & Sedat, J.W. "T⁵M: 3D widefield light microscopy with better than 100 nm axial resolution" *J. Microsc.* **195**, 10-16 (1999).
5. Klar, T.A., Jakobs, S., Dyba, M., Egnér, A. & Hell, S.W. "Fluorescence microscopy with diffraction resolution barrier broken by stimulated emission" *PNAS* **97**, 8206-8210 (2000).
6. Kawata, S. "The optical computed tomography microscope" in *Advances in Optical and Electron Microscopy*, Vol. 14, T. Mulvey & C.R.J. Sheppard, eds., (Academic Press Limited, San Diego, 1994).
7. Kikuchi, S., Sonobe, K., Sidharta, L.S. & Ohya, N. "Three-dimensional computed tomography for optical microscopes" *Opt. Commun.* **107**, 432-444 (1994).
8. Sharpe, J. et al. "Optical projection tomography as a tool for 3D microscopy and gene expression studies" *Science* **296**, 541-545 (2002).
9. Wilke, A.L., Jordan, S.A., Sharpe, J.A., Price, D.J. & Jackson, I.J. "Widespread tangential dispersion and extensive cell death during early neurogenesis in the mouse neocortex" *Dev. Biol.* **267**, 109-118 (2004).
10. Shaw, P.J., Agard, D.A., Hiraoka, Y. & Sedat, J.W. "Tilted view reconstruction in optical microscopy" *Biophys. J.* **55**, 101-110 (1989).
11. Bradl, J., Hausmann, M., Ehemann, V., Komitowski, D. & Cremer, C. "A tilting device for three-dimensional microscopy: application to in situ imaging of interphase cell nuclei" *J. Microsc.* **168**, 47-57 (1992).

12. Cogswell, C.J., Larkin, K.G. & Klemm, H.U. "Fluorescence microtomography: multi-angle image acquisition and 3D digital reconstruction" *SPIE Proc.* **2655**, 109-115 (1996).
 13. Kozubek, M. et al. "Automated microaxial tomography of cell nuclei after specific labeling by fluorescence in situ hybridisation" *Micron* **33**, 655-665 (2002).
 14. Bradl, J., Hausmann, M., Schneider, B., Rinke, B. & Cremer, C. "A versatile 2π -tilting device for fluorescence microscopes" *J. Microsc.* **176**, 211-221 (1994).
 15. Verveer, P.J. & Jovin, T.M. "Improved resolution from multiple images of a single object: application to fluorescence microscopy" *Appl. Opt.* **37**, 6240-6246 (1998).
 16. Huisken, J., Swoger, J., Del Bene, F., Wittbrodt, J. & Stelzer, E.H.K. "Optical sectioning deep inside live embryos by selective plane illumination microscopy" *Science* **305**, 1007-1009 (2004).
 17. Voie, A.H., "Imaging the intact guinea pig tympanic bulla by orthogonal-plane fluorescence optical sectioning microscopy", *Hearing Research* **171**, 119-128 (2002).
 18. Dodt, H.-U., Leischner, U., Schierloh, A., Jährling, N., Mauch, C.P., Deininger, K., Deussing, J.M., Eder, M., Zieglgänsberger, W. & Becker, K., "Ultramicroscopy: three-dimensional visualization of neuronal networks in the whole mouse brain" *Nat. Methods* **4**, 331-336 (2007).
 19. Verveer, P.J., Swoger, J., Pampaloni, F., Greger, K., Marcello, M., & Stelzer, E.H.K., "High-resolution three-dimensional imaging of large specimens with light sheet-based microscopy" *Nat. Methods*, **4**, 311-313 (2007).
 20. Greger, K., Swoger, J., & Stelzer, E.H.K., "Basic building units and properties of a fluorescence single plane illumination microscope" *Rev. Sci. Instrum.* **78**, 023705 (2007).
 21. Abbott, A. "Biology's new dimension" *Nature* **424**, 870-872 (2003).
 22. Born, M. & Wolf, E. *Principles of optics*, 7th Ed. Cambridge University Press, Cambridge, CB2 2RU, U.K., 1999.
 23. Sätzler, K. & Eils, R. "Resolution improvement by 3-D reconstructions from tilted views in axial tomography and confocal theta microscopy" *Bioimaging* **5**, 171-182 (1997).
 24. Kikuchi, S., Sonobe, K. & Ohya, N., "Three-dimensional microscopic computed tomography based on generalized Radon transform for optical imaging systems" *Opt. Commun.* **123**, 725-733 (1996).
 25. Brooks, R.A. & Di Chiro, G. "Theory of image reconstruction in computed tomography" *Radiology* **117**, 561-572 (1975).
 26. Frasch, M. "The maternally expressed *Drosophila* gene encoding the chromatin-binding protein BJI is a homolog of the vertebrate gene Regulator of Chromatin Condensation, RCC1" *EMBO J.* **10**, 1225-1236 (1991).
 27. Gortchakov, A.A. et al. "Chriz, a chromodomain protein specific for the interbands of *Drosophila melanogaster* polytene chromosomes" *Chromosoma* **114**, 54-65 (2005).
 28. See the Medaka Expression Pattern Database (MEPD), <http://pubserv1.embl.de:8280/pubserv/servlet/de.embl.th.mepd.servlets.MdbShowClone01?cloneID=1251>.
 29. Qiring, R. et al. "Large-scale expression screening by automated whole-mount in situ hybridization" *Mech. Dev.* **121**, 971-976 (2004).
 30. Engelbrecht, C.J. & Stelzer, E.H.K. "Resolution enhancement in a light-sheet-based microscope (SPIM)" *Opt. Lett.* **31**, 1477-1479 (2006).
 31. Swoger, J., Huisken, J. & Stelzer, E.H.K. "Multiple imaging axis microscopy improves resolution for thick-sample applications" *Opt. Lett.* **28**, 1654-1656 (2003).
 32. Brown, L.G. "A survey of image registration techniques" *ACM Computing Surveys* **24**, 325-376 (1992).
 33. Press, W.H., Teukolsky, S.A., Vetterling, W.T., & Flannery, B.P. *Numerical Recipes in C++*, 2nd Ed. (Cambridge University Press, Cambridge, U.K., 2002).
-

1. Introduction

1.1. Techniques for high-resolution optical imaging in biology

A serious problem that is encountered in traditional optical microscopy is the fact that the axial resolution in a three-dimensional (3D) image stack is not as good as the lateral resolution. This is a limitation for modern biological applications, where the axial resolution (typically $> \sim 1 \mu\text{m}$) is insufficient to resolve sub-cellular phenomena. Several novel types of instrumentation, such as confocal 4Pi [1] and theta [2] microscopies, optical coherence tomography (OCT) [3], I³M [4], and stimulated emission depletion (STED) microscopy [5], have been proposed to overcome this problem. Although these techniques are effective in the applications for which they were designed, they are all technologically challenging. Apart from theta microscopy, all are based around hardware solutions involving coherent optics, which places strict requirements on the hardware's precision and stability.

Another approach which does not require coherent optics is based on the extension of X-ray computed tomography (CT) techniques into the visible region of the spectrum. Although the details vary, all such techniques involve combining information obtained from multiple views of the sample. Initial implementations involved computation of multiple projections from 3D optical image stacks [6,7], which were then combined by traditional CT processing. More recently optical projection tomography (OPT) [8], in which the projections are generated directly by the optical system, has been shown to be a valuable tool for developmental biology [9].

Unlike in X-ray CT, in which all images are projections, in optical microscopy it is possible to generate 3D optical image stacks from multiple viewpoints. Such multi-view stacks have been generated by both traditional wide-field (i.e. non-sectioned) [10-13] and confocal [14] methods. Although the processing required to do tomography with true 3D data stacks is somewhat different from classical CT (see Section 1.3, below), such techniques generally require fewer views of the sample to achieve optimal image quality.

A generalization of the idea of multi-view image fusion is given in Verwee & Jovin [15], in which images recorded with different *modes* of microscopy (in their case, confocal and wide-field), rather than different *viewing directions*, are combined. This method of multi-mode image fusion is equally applicable to multi-view image fusion, as is demonstrated using the MVD-MAPGG algorithm in the Results section, below.

Light sheet based fluorescence microscopy techniques, including such variants as Selective (or Single) Plane Illumination Microscopy (SPIM) [16], Orthogonal Plane Fluorescence Optical Sectioning (OPFOS) [17], and ultramicroscopy [18], make it possible to generate 3D optically sectioned fluorescence images of biological samples. If, as is the case with SPIM, multiple 3D stacks are acquired from different viewing angles, this data is particularly well suited to further enhancement by multi-view image fusion. This is because it provides quantitative 3D fluorescence data with a high signal-to-noise ratio (SNR) that permits the use of non-linear reconstruction algorithms. Recently we presented some results from a SPIM system, including our initial multi-view reconstructions [16,19]. In this work we describe the details of and improvements to the image processing algorithms used to combine data from multiple viewing directions, and demonstrate their effectiveness on several biological SPIM data sets.

A schematic of the SPIM setup is shown in Fig. 1. The detection arm (OL, EM, TL, and CCD) constitutes a traditional wide-field fluorescence microscope. Illumination is achieved through a cylindrical lens (CL), which forms a thin light sheet whose centre coincides with the focal plane of the detection optics inside the sample (S). This configuration provides optical sectioning and reduced photo-damage to the sample, because only the plane on which the CCD camera is focused is illuminated. The sample is typically mounted in a cylinder of agarose gel immersed in an appropriate aqueous medium. This allows living samples to be maintained in a viable environment while minimizing aberrations due to refractive index mismatches. The gel containing the sample is attached to a stage that translates the sample along the detection axis, so that a 3D image stack of the sample can be recorded. It also allows rotation of the sample around the vertical axis (perpendicular to both the illumination and detection axes), so that multiple 3D stacks can be generated along different viewing angles. Further details of the SPIM hardware can be found in Greger *et al* [20].

1.2. Multi-view image processing goals

For many applications (e.g. thin cells cultured on a flat cover slip) single-view images of the sample are adequate. However, in recent years there has been a growing interest in quantitative 3D imaging of thicker specimens [21], and such imaging presents challenges that are not issues in traditional 2D microscopy.

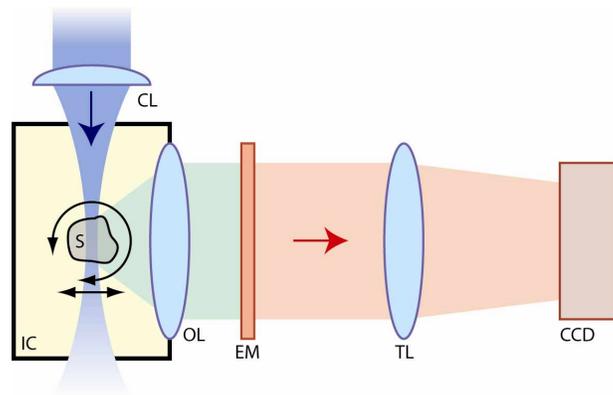


Fig. 1. SPIM schematic. The sample (S) is illuminated by a thin light sheet generated by passing a collimated laser beam through a cylindrical lens (CL). The region of the sample that is imaged onto the CCD camera by the objective (OL) and tube (TL) lenses is illuminated by the light sheet. The emission filter (EM) blocks scattered illumination light. The sample is scanned along the detection axis to create a 3D image stack. It can then be physically rotated to different orientations and re-scanned, generating sets of 3D stacks along different viewing angles. An immersion-medium-filled chamber (IC) encloses the specimen and reduces the negative effects of aberration-inducing interfaces.

For simplicity we restrict ourselves in what follows to a discussion of fluorescent samples, although similar arguments could be made for other contrast mechanisms.

There are two main categories into which we divide biological samples: (nearly) transparent and (partially) opaque. By transparent samples we mean those with relatively low scattering/absorption, so that the image resolution does not vary significantly with depth in the sample. For the sample fluorophore distribution sketched in Fig. 2(a), the images detected along two perpendicular axes are shown in Fig. 2(b) and (c). In such cases image distortions are primarily caused by the optical system, i.e. the image is blurred by the point-spread function (PSF) of the microscope. These images have an anisotropic resolution (most optical microscopes have poorer resolution axially than laterally), but the resolution is spatially invariant (the amount of blurring does not depend on the location within the sample) [22]. Since each view has its best resolution along a different direction (e.g. vertical and horizontal in Figs. 2(b) & (c), respectively), the object of the image fusion is to achieve isotropic resolution by combining the best information from each view (as in Fig. 2(d)).

The case of a partially opaque sample is illustrated in the bottom row of Fig. 2. Here the primary problem is not anisotropic resolution, but that the image quality is degraded as one images deeper inside the sample. In this sort of sample, the resolution will also be anisotropic (when imaged with a traditional single-objective microscope), but when the absorption is high this anisotropy is often a relatively minor effect. Multi-view imaging allows us to observe different regions of the sample (Figs. 2(f) and (g)). In this case the goal of multi-view fusion is to combine the information about the different regions of the sample, as illustrated in Fig. 2(h).

These two categories of samples are extreme cases: many samples of practical interest lie somewhere in between and present characteristics of both types. The fusion process described in this work can be beneficial for them in both of these respects. It should be noted that previous work in multi-view microscopy has concentrated primarily on resolution improvement in relatively transparent samples (e.g. Sätzler & Eils [23], Kikuchi *et al.* [24]). Two useful properties of the multi-view fusion algorithm described in this work are therefore that it is robust enough to handle samples that lie at or anywhere between these two extremes, and that the amount or degree of each type of degradation need not be specified.

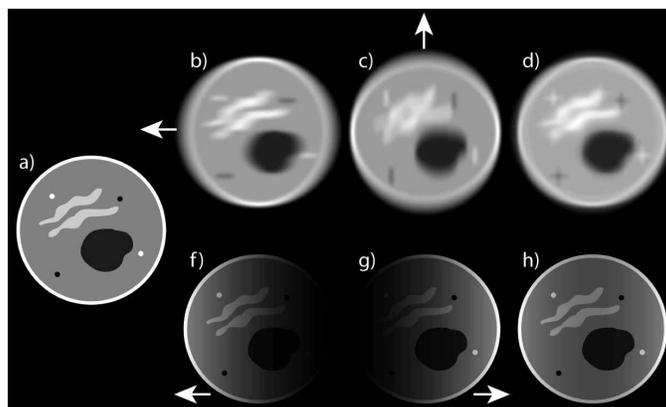


Fig. 2. Illustration of image distortions caused by anisotropic and spatially varying image quality, and the compensation thereof by fusing two images. a) The undistorted sample fluorophore distribution. Top row: Images distorted by reduced resolution horizontally (b) or vertically (c). d) Fusion of b) and c) shows improved resolution isotropy. Bottom row: Images distorted by absorption while detecting from the left (f) and from the right (g). h) Fusion of f) and g) shows improved spatial coverage of the sample. Arrows indicate the directions along which the light is detected.

2. Results

2.1 Transparent samples

2.1.1 Types of image fusion

The first example of multi-view fusion that we present, an example of a (mostly) transparent specimen, uses images of a grain of paper mulberry pollen. Thirty-six 3D stacks with equal angular spacing were taken of the autofluorescence intensity distribution of the pollen grain. From one of these 36 views we extracted optical slices through the middle of the pollen grain, in a plane parallel to both the illumination & detection axes (Fig. 3(a)). Despite the optical sectioning provided by the SPIM light sheet illumination, the inferior resolution along the detection axis is evident. Figure 3(b) shows the corresponding slice through a data set taken in an orientation perpendicular to that in Fig. 3(a), in which the inferior resolution axis is orthogonal to that in Fig. 3(a).

The arithmetic-mean fusion of the 36 views is shown in Fig. 3(c). The resolution is clearly more isotropic than in the single views, and the fact that the grain consists of a spherical shell (the sporoderm) surrounding an inner core becomes apparent. However, fine details of the structure are still obscured. Figure 3(d) shows the result of image fusion by weighted spectral averaging (similar to the Fourier reconstruction used in CT [25]). The contrast and sharpness of the slices is slightly better than that obtained by the arithmetic mean, which is a result of the preferential selection of high-signal components in the weighted spectral fusion.

The results of the multi-view deconvolution (MVD) fusions using iterative Wiener (MVD-Wiener) and maximum *a posteriori* with Gaussian noise and prior [15] (MVD-MAPGG) methods (see Methods for details) are shown in Figs. 3(e) and (f), respectively. Both of these methods make use of the PSF to correct for the response of the optical system, and both are a significant improvement over either of the non-iterative fusions: the structures in the interior are much better resolved, and it is clear that the grain is encased in a continuous thin shell. The results of the two iterative fusions are very similar, although the MVD-Wiener does tend to exhibit slightly higher background and ringing artifacts.

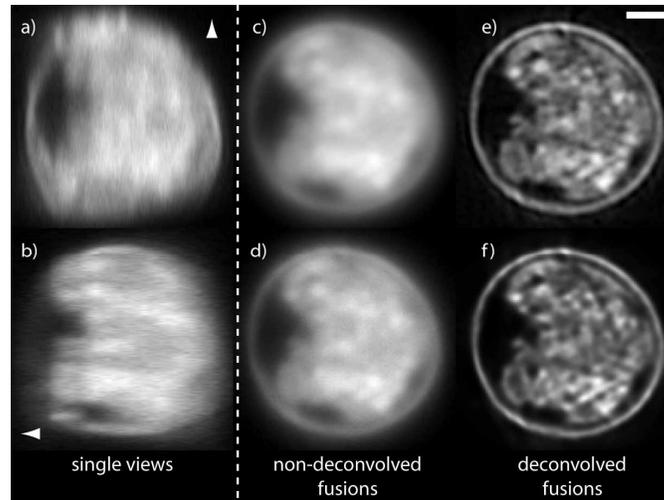


Fig. 3. Slices perpendicular to the rotation axis, through 3D autofluorescence data stacks of a paper mulberry pollen grain. a) 0°-view. b) 90°-view. c-f) Fusions of 36 views by: c) arithmetic mean; d) weighted spectral mean; e) MVD-Wiener; f) MVD-MAPGG. Arrowheads in a) and b) indicate the directions of the respective detection axes. Scale bar = 3 μ m. See also Movie 1 (1.55MB), slices rotating through the grain of paper mulberry pollen, for which the left image is the 0° view, and the right is the MVD-MAPGG fusion.

Overall, there is a clear progression from the raw data sets to the non-iterative fusions to the MVD fusions. This is a result of the increased information input that is involved in the fusions: those that do not involve deconvolution are improved over the raw data because they combine information from different (anisotropic) views. In addition to this, the MVD fusions also make use of the PSFs, which contain information about the optical system used to generate the raw images.

2.1.2 Effects of the number of views fused

One question that arises when doing multi-view fusions is: What is the optimal number of views to combine? A large number of views will clearly contain more (possibly redundant) information about the sample, but will also require extended data acquisition and processing times, and can result in increased photo-damage to the sample. We illustrate the effect of the number of views on the fusion results in Fig. 4, where MVD-MAPGG fusions have been done for sets of 1, 2, 4, 6, 12, and 18 views of a pollen grain similar to that used in Fig. 3. As can be seen by comparing Figs. 4(a) & (b), combining views along the same axis does not improve the image isotropy, although two anti-parallel views are not identical because even the small pollen grain is not completely transparent. As the number of views is increased, (Figs. 4(c)-(f)), the anisotropy-related artifacts in the images are reduced. This effect can be observed in the smoothness of the shell of the pollen grain in the top row, as the number of views increases. It is also evident in the increased radial symmetry of the corresponding spectral plots, where the “gaps” between the “spokes” in the spectra are reduced as more views are included (the “gaps” are regions in which the reconstruction contains little or no information about the sample).

For this sample, a combination of 18 views (with consecutive views separated by 20°) provides isotropic resolution, and the inclusion of more views does not significantly improve the image quality. This data set was measured using SPIM with a water-immersion objective lens with a numerical aperture (NA) of 0.9, which has a detection half-angle of $\alpha = 42.6^\circ$ (where $NA = n_i \sin(\alpha)$, $n_i = 1.33$ is the refractive index of water). It is quite difficult in general to estimate *a priori* the optimal number of views to process, because this number

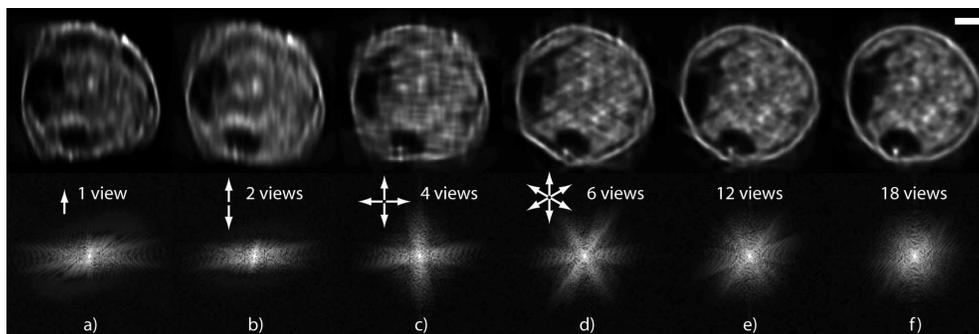


Fig. 4. Effects of the number of views used in the MVD-MAPGG fusion of paper mulberry pollen autofluorescence images. Top: slices perpendicular to SPIM rotation axis, for different numbers of fused images. Bottom: corresponding power spectra (plotted with a non-linear look-up table to emphasize the high-frequency components). The detection axes are indicated by the white arrows in a-d. Scale bar = 3 μ m.

will depend not only on the configuration of the instrument that acquires the data (e.g. the NA of the objective lens, and, for a SPIM system the light sheet thickness), but also on the sample properties. The amount of optical attenuation (i.e. the scattering and/or absorption properties) and the physical structure of the sample will affect which regions will be well imaged from a given viewing direction, and therefore the optimal number and arrangement of the views. However, we have found that for mostly transparent samples such as the pollen grain in Fig. 4, as a rule of thumb the optimal number of views can be estimated using $N_{optimal} \sim 360^\circ/(\alpha/2)$, which results in $N_{optimal} \sim 17$ for the specimen shown in Fig. 4.

2.2 Opaque samples

2.2.1 Sub-cellular resolution over an entire embryo

In addition to nearly transparent samples such as the pollen grain shown in the previous section, multi-view image fusion is useful for improving data sets from thick, partially opaque specimens [16]. A *Drosophila melanogaster* embryo, which is difficult to image in its entirety because of its highly scattering tissue, is shown in Fig. 5. The nuclear proteins BJ1 [26] (green) and Chriz [27] (red) are visualized. Figure 5(a) shows projections through a single-view data stack, in which less than a quarter of the specimen is visible. In contrast, the entire embryo is clearly visible in the 8-view MVD-Wiener fusion shown in Fig. 5(b). Figure 5(c) shows the sub-region of the fused data set indicated in Fig. 5(b), which clearly shows the individual nuclei, and the separate domains of the two labeled proteins. Even though a low power objective lens ($M = 10\times$, $NA = 0.3$ water immersion) was used to acquire the raw data, the multi-view fusion yields sub-cellular resolution over the entire labeled outer region of the embryo.

2.2.2 Imaging deep within an embryo

Figure 5 demonstrates whole-embryo multi-view image fusion. However, because of the sample properties only the outer layers contain visible fluorescent structures. The sample therefore does not give a clear indication of the 3D depth penetration achievable with multi-view SPIM data. To demonstrate that this technique can provide improved resolution within a thick biological specimen, we show images of a stage 32 Medaka fish embryo with a general nuclear staining (green) and *in situ* hybridization for the gene McF0001MGR-1G19bd1 (red) [28]. Figures 6(a) and (b) depict maximum-value projections along two orthogonal axes of the embryo. Images in the left column are from a single-view data set; those on the right are from a 6-view MVD-MAPGG fusion. The improvements in sharpness, contrast, and uniformity of the fused images are apparent. The inset shows a traditional wide-field

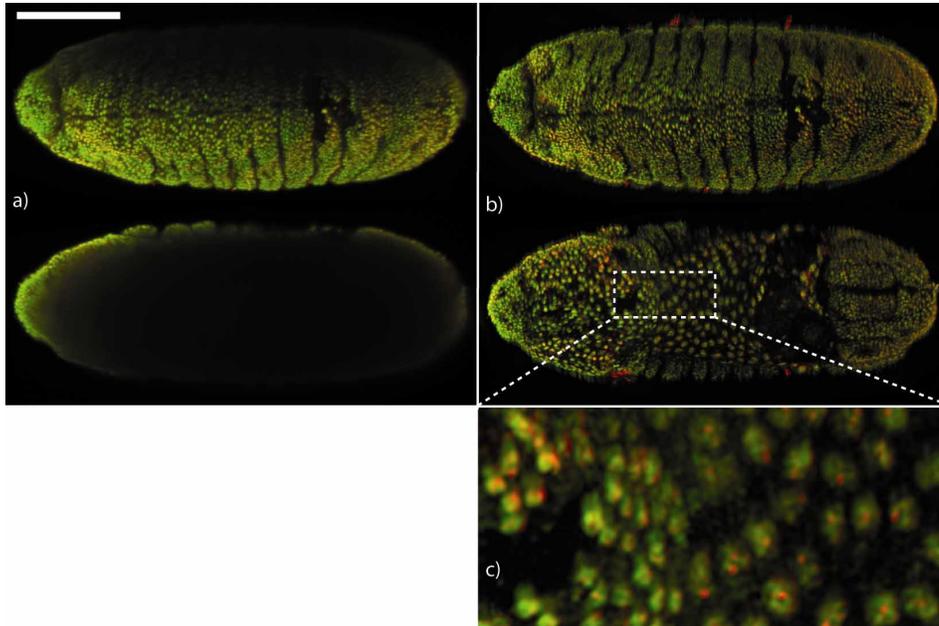


Fig. 5. *Drosophila melanogaster* embryo. Green: uncondensed DNA; red: condensed DNA. a) Maximum-value projections of a single-view data set, through the ventral (top) and dorsal (bottom) halves of the embryo. b) as a), but for an 8-view MVD-Wiener fusion. c) Expanded image of the 100 μm wide region indicated in b). Scale bar = 100 μm . See also Movie 2, maximum-value projections through the MVD-Wiener fusion of the *Drosophila* embryo.

transmission image, in which the *in situ* localization is in agreement with that seen in the SPIM images: McF0001MGR-1G19bd1 is primarily expressed in the tectum proliferation zone, the pineal gland, and the telencephalon.

To provide a better visualization of the 3D imaging capability of multi-view image fusion, Fig. 6(c) shows a slice through the head of the embryo at the position indicated by the dashed line in Fig. 6(a). Much of the detail in the single-view slice (left) is obscured by blurring and absorption. In contrast, the fusion of six views (right) shows considerable detail throughout the cross-section of the embryo. Various layers in the eye can be easily identified, as well as the structures of other tissues within the head of the embryo. (Note that the intensity of the red channel in Fig. 6(c) has been increased relative to that in Figs. 6(a) and (b) to make the expression of McF0001MGR-1G19bd1 in the forebrain and the lens of the eye apparent.)

In the single-view images dark stripes can be seen running horizontally through the embryo in Fig. 6. These are shadows caused by strong absorption of the illumination light, which can occur in all types of fluorescence microscopy. However, because of the collimated nature of SPIM illumination these artifacts are more obvious, and can extend through the entire sample. This can make quantification of fluorophore distributions from single-view SPIM data sets difficult, if the sample exhibits significant amounts of pigmentation. However, in addition to the overall improvement in image resolution and uniformity provided by the six-view MVD-MAPGG fusion, this data also demonstrates that multi-view fusion of SPIM data sets effectively reduces shadowing artifacts.

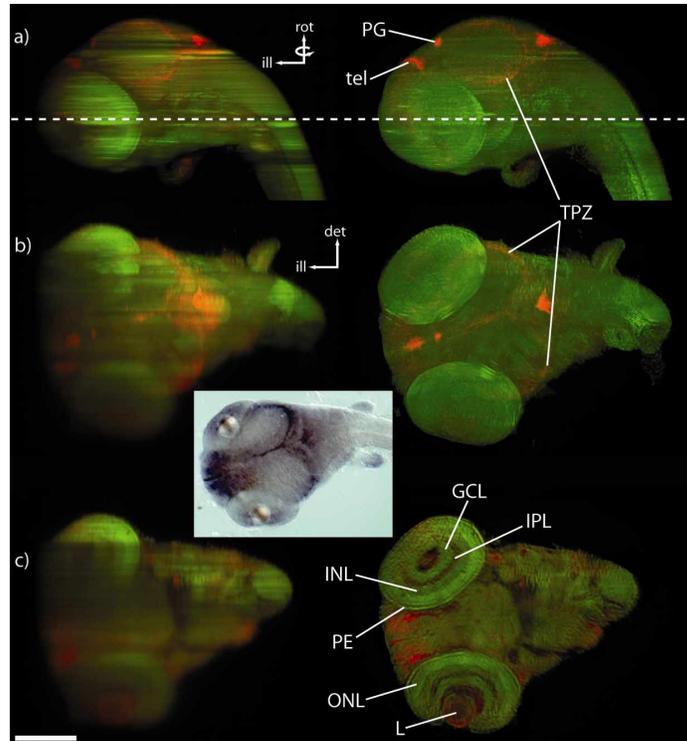


Fig. 6. Medaka embryo, stage 32, nuclear label (green) and Mcf0001MGR-1G19bd1 *in situ* hybridization (red). Left: single-view images; right: 6-view MAPGG fusion. a,b) Maximum-value projections along orthogonal axes. The regions of expression of Mcf0001MGR-1G19bd1 in the tectum proliferative zone (TPZ), pineal gland (PG), and telencephalon (tel) are clearly visible. For comparison, the inset shows the traditional blue-labeled transmission image. c) Slice at the depth indicated by the dashed line in a). Internal structures such as the lens (L), pigmented epithelium (PE), ganglion cell layer (GCL), outer (ONL) and inner nuclear layers (INL), and the inner plexiform layer (IPL) of the retina are well-defined in the fusion. The illumination (ill), detection (det), and rotation (rot) axes are indicated for the single-view images. Scale bar = 200 μ m. See also Movie 3, an animation of the MVD-MAPGG fusion showing maximum-value projections at the top and slices at the position indicated by the white line at the bottom.

3. Methods

3.1. Sample preparation

Pollen: Paper mulberry pollen grains (Polysciences Inc., Warrington, PA, USA) were used as supplied.

Drosophila: Fixed blastocyte samples were labeled for the *Drosophila* proteins BJ1 [27] (FITC labeled GAM Ig(H+L), Jackson Imm. Res.; monoclonal mouse antibody) and Chriz [28] (TRITC labeled GAR Ig(H+L), Jackson Imm. Res.; rabbit polyclonal antibody).

Medaka: Whole-mount *in situ* hybridization was performed using digoxigenin-labeled RNA riboprobes as described previously in Quiring *et al* [29]. After *in situ* hybridization, embryos were incubated 10 minutes in 1 mM Sytox green (diluted in 1 M Tris, pH 7.5) and washed three times with 1 M Tris, pH 7.5.

All samples were mounted in 1% (or slightly lower) low melting agarose gel (Sigma-Aldrich, St. Louis, MO, USA) prepared using a phosphate buffered saline (PBS) solution.

3.2. Image acquisition

The details of the SPIM are given in [16,20], and sketched in Fig. 1. Briefly, after mounting in the agarose gel the specimens were attached to the translation stage of the SPIM and stacks along different viewing directions were recorded sequentially. Detection was performed using water dipping microscope objective lenses ranging from 63×/0.9 for the pollen ($\lambda_{\text{ill}} = 488 \text{ nm}$, $\lambda_{\text{det}} > 510 \text{ nm}$) to 10×/0.3 for the *Drosophila* embryo ($\lambda_{\text{ill},1} = 488 \text{ nm}$, $\lambda_{\text{det},1} > 510 \text{ nm}$, $\lambda_{\text{ill},2} = 543 \text{ nm}$, $570 \text{ nm} < \lambda_{\text{det},2} < 650 \text{ nm}$). The exception was the Medaka embryo, which was imaged using a 5×/0.25 air objective ($\lambda_{\text{ill},1} = 488 \text{ nm}$, $\lambda_{\text{det},1} > 510 \text{ nm}$, $\lambda_{\text{ill},2} = 543 \text{ nm}$, $570 \text{ nm} < \lambda_{\text{det},2} < 650 \text{ nm}$). The camera used was a 12-bit Hamamatsu Orca-ER with a pixel pitch of 6.45 μm . The technical details of the optimization of the illumination and detection systems are given in Engelbrecht [30]; as described therein, for all imaging in the present work, the SPIM illumination light sheet was configured so as to illuminate the field of view of the CCD camera with $\leq 1 - 2^{-0.5} \sim 29\%$ variation in intensity.

The deconvolution methods used in this work are non-blind; i.e. they require as inputs the PSF of each view used. These were generated by imaging sub-resolution fluorescent beads by SPIM under the same conditions used for the samples to be deconvolved. Several (6-10) beads were selected, registered, and averaged to reduce noise, producing an estimate of the system PSF. To generate the appropriate PSF for each view, multiple copies of the original system PSF are made and pre-processed as outlined in Section 3.3.1, below.

3.3. Image processing

In general the image processing algorithm used in this work can be divided into three main steps: *pre-processing* (consisting of axial re-sampling and stack rotation), *registration* of the individual views, and the *fusion* of the registered views into a single description of the sample. The details of the algorithm used for processing 3D multi-view data sets are given in Fig. 7. The technique is an adaptation (to the data SPIM data sets generated by rotating the sample around the vertical axis) and extension (by including the MVD options) of that described in Swoger *et al.* [31] for the four-view fusion of stacks acquired with a specific (tetrahedral) geometry.

3.3.1 Pre-processing

The purpose of pre-processing the image data is to transform the data sets so that the subsequent step of registration is simplified. Proper pre-processing of SPIM data, for example, makes it possible to register the images by a 3D translation, rather than requiring a 12-parameter affine transformation. The pre-processing transformation is based on known parameters of the imaging system (the pixel pitch, optical magnification, angle of rotation between the views, and the spacing of the slices in the stacks), rather than information obtained from the images themselves. As illustrated in Fig. 7, each stack is rescaled axially by bicubic interpolation so that the axial slice spacing is equal to the lateral pixel spacing. After this the 3D stacks are rotated (again, using bicubic interpolation) so that they share a common orientation (generally, but not necessarily, the orientation of the first stack).

3.3.2 Registration

Image registration is the process of aligning multiple images so that the common features overlap. [32] There are two main questions that need to be answered when choosing a registration method: 1) what type of transformation is required to perform the registration, and 2) how are the parameters for that transformation to be determined. Transformations ranging from simple translations to higher order global or local transformations can be applied, depending on the properties of the data sets. In general the higher-order transformations can correct for more complicated types of mis-registration (e.g. resulting

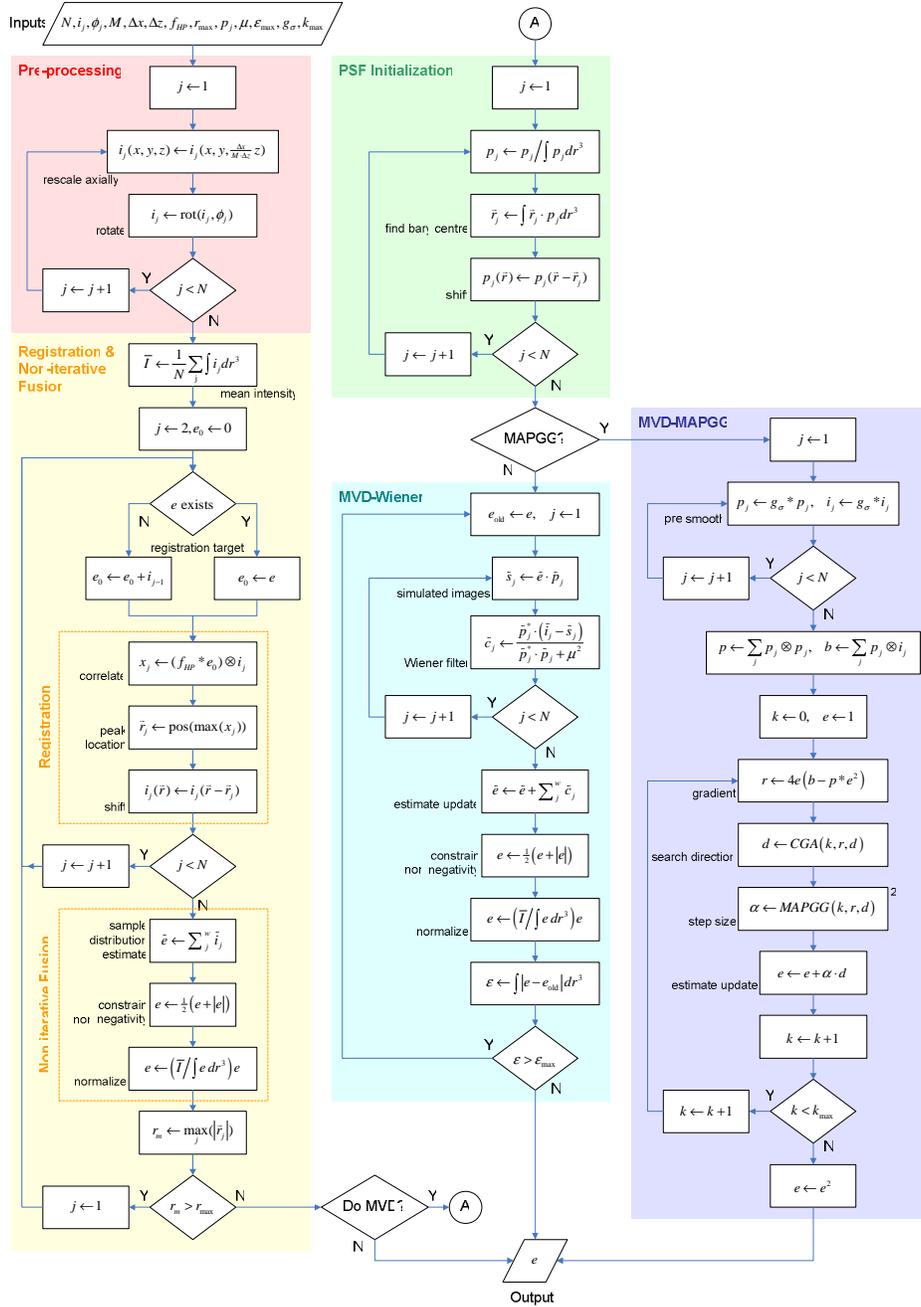


Fig. 7. Image processing algorithm. **Inputs:** N , number of views; i_j , j^{th} image; ϕ_j , j^{th} viewing angle; M , optical magnification; Δx , pixel pitch (lateral); Δz , slice spacing (axial); f_{HP} , high-pass spatial frequency filter; r_{max} , maximum registration error; p_j , j^{th} PSF; μ , MVD-Wiener regularization parameter; ϵ_{max} , MVD-Wiener termination parameter; g_σ , MVD-MAPGG Gaussian filter; k_{max} , MVD-MAPGG maximum number of iterations. **Output:** e , sample distribution estimate. **Internal:** $\bar{r} = (x, y, z)$, position; \bar{s} , spatial frequency; \bar{I} , mean image integral; e_0 , registration target; x_j , j^{th} cross-correlation; \bar{r}_j , position of peak of x_j ; r , gradient vector; d , search direction; α , step size along d . **Functional:** \sum_j^w , weighted average (see text); CGA, conjugate gradient algorithm [33]; MAPGG, maximum *a posteriori* with Gaussian noise and prior algorithm [15]; *, convolution operator; \otimes , correlation operator; \tilde{a} , Fourier transform of a .

from inaccuracies in the rotation stage or in the calibration of the image scaling, or those due to optical aberrations in the imaging system), but they are difficult to automate and computationally time-consuming, and are therefore to be avoided if possible.

The image registration employed in this work is purely translational, which is advantageous because it is a relatively simple transformation. The shift vectors required for the translations are found by cross-correlation, implemented using fast Fourier transform algorithms. In principle these shift vectors might be extracted directly from the parameters sent to the scanning hardware during image acquisition, but this option was not included in our implementation of the algorithm.

We have found that in some samples (the “mostly opaque” ones), registration of all of the views to a fixed target often fails because some views have virtually no overlap with the target. To overcome this difficulty, in the initial stage of registration we have implemented a “running sum” as the target; e.g. if one has views at 0°, 90°, and 180°, the 0° is used as the target to which the 90° view is registered, and then the sum of the 0° and (registered) 90° views is used as the target for the 180° registration. The spatially non-uniform image quality in the single view of the embryo in Fig. 5(a) is apparent. Figure 8 shows cross-sections from the same data set, which illustrate the need for the running-mean correlation target described above.

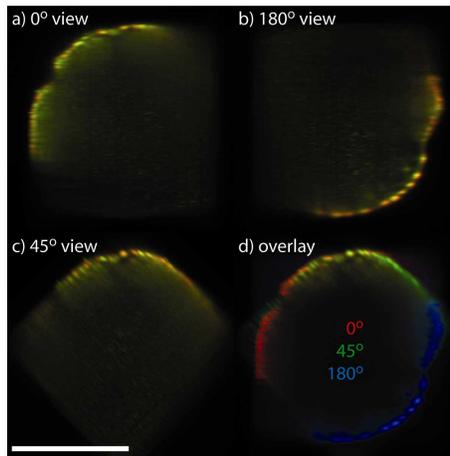


Fig. 8. Slices perpendicular to the anterior/posterior axis of the *Drosophila melanogaster* embryo shown in Fig. 5. Slices from a) a view showing the upper left portion of the embryo clearly; b) a view taken at 180° with respect to the orientation in a); c) a view taken at 45° with respect to a). d) Overlay of the slices in a-c). The views in a) and b) contain very little overlapping information, which would make them difficult to register in isolation. Scale bar = 100 μm. See also Movie 4, an animation of progressive construction of a running-mean registration target. Individual cross-sections are shown in the red channel, and the green channel shows the running mean registration target as each view is successively added.

After the initial round of registration, an initial fusion of the images is generated by taking the weighted average of the registered views in Fourier space (see Fig. 7)

$$\Sigma_j^w (\tilde{i}_j) = \Sigma_j (w_j \cdot \tilde{i}_j),$$

where \tilde{i}_j are the data sets to be averaged and w_j the corresponding weights. We assume that the image noise is governed by Poisson statistics, so that the weighting of the average is by the expected signal-to-noise ratio (SNR):

$$w_j = \frac{\sqrt{|\tilde{i}_j|}}{\sum_j (\sqrt{|\tilde{i}_j|})}$$

This fusion is then used for a second registration stage, in which the target is the fusion from the first stage (rather than a running sum of the individual views). The process is repeated until the registration corrections become small (typically less than 1 pixel), which generally occurs within at most three iterations. We have found that this iterative procedure, although computationally time-consuming, generally provides a small but significant improvement in the quality of the final reconstruction (data not shown).

3.3.3 Fusion

At this point it is possible to halt the processing; the resulting output is the registered image stacks and the initial, non-iterative fusion. However, as is shown in Fig. 3, a significant further improvement to the fusion can often be gained by fusing the registered images using a multi-view deconvolution (MVD). This requires that the PSFs of the views are known, and that the images are of sufficiently high SNR to allow the deconvolution to be effective. The iterative deconvolution steps are outlined on the right in Fig. 7 [33]. First the PSFs are normalized to an integral of unity and shifted so that they are centered at the origin of the coordinate system. This shift is required so that the deconvolution process does not introduce unwanted shifts which would effectively “unregister” the images.

We have investigated two methods for performing the MVD: an iterated Wiener filter approach (MVD-Wiener), and a MAPGG algorithm (MVD-MAPGG). The MVD-MAPGG algorithm is summarized in Fig. 7 and is an adaptation of previously described work [15]. The MVD-Wiener process (also outlined in Fig. 7) involves creating a set of “simulated images”, \tilde{s}_j , by convolving the initial sample estimate, e , with the PSFs. A set of correction terms, \tilde{c}_j , are then generated from these using a regularized inverse filter. These correction terms represent a measure of how close the simulated images (what we would expect the images to look like, given the estimated sample fluorophore distribution, e) are to the actual, measured images. These correction terms are then combined by the weighted average described above, and added to the estimate from the previous iteration to yield an updated estimate. After constraining the sample distribution to be non-negative (for fluorescent imaging) and normalization, the new estimate is compared with the previous one. If the change is sufficiently small the fusion is complete; if not, another iteration of the deconvolution is performed. The regularization parameter μ was optimized manually by pre-deconvolving a small region of interest in a single view of each data set, and then applying this value in the full MVD. This type of multi-view deconvolution has recently been shown to allow SPIM to exceed the resolution of confocal microscopy, even when the confocal uses lenses with $NA > 1.2$ [22].

The algorithm (except for the MDV-MAPGG routine) was implemented in MATLAB version 7.0.4 (The MathWorks, Inc.). However, to improve processing speed some of the calculation-intensive modules (e.g. Fourier transforms) are implemented using Visual C# .NET (Microsoft Corp.) and call Intel Integrated Performance Primitives version 4.1 (Intel Corp.). Although the algorithm is not currently implemented such that it can process large multi-view data stacks in real time, it is fast enough to be of use for many non-real-time applications. For the data presented in this work, calculation times varied from a few hours (for the pollen sample, Figs. 3-4, where each view was $256 \times 256 \times 256$ voxels) to over-night (for the Medaka embryo, Fig. 6 with $609 \times 450 \times 788$ voxels) on a desktop PC with a Xeon 2.8GHz CPU. It is thus possible to fuse time-lapse series of multi-view data sets in order to follow, e.g. organogenesis during embryo development, although such calculations will run over several days.

The MAPPG algorithm was adapted from Verveer & Jovin [15], by allowing more than two input images to be used, and by allowing large data sets to be divided into overlapping blocks, which could be processed separately and subsequently stitched together. This portion of the algorithm was implemented using the programming language Python (www.python.org) with the Numarray extension (www.stsci.edu/resources/software_hardware/numarray) for numerical computing. The algorithm was executed on a cluster of 10 computers with dual-Pentium 4 2.4 GHz processors and 2GB of memory each, running the Linux operating system with the OpenMosix clustering extension. Execution times for the MAPGG routine varied depending on the data sets, but are usually on the order of an hour provided the calculation was properly distributed over the cluster. The source code used to implement the algorithms described in this manuscript is available on request (contact E.H.K. Stelzer, e-mail stelzer@embl.de).

4. Discussion

The method described here for processing multi-view data has been shown to be advantageous under a range of conditions. For samples that can be imaged in their entirety in all views (e.g. the pollen grain in Figs. 3 & 4), the fusion provides isotropic resolution that can be higher than that in any of the unprocessed individual views. At the other extreme, when each view shows only a portion of the specimen (e.g. the *Drosophila* embryo in Fig. 5) the fusion algorithm can combine these results in a “smooth mosaic” covering the entire sample volume. Partially transparent samples such as the Medaka embryo in Fig. 6 can benefit from both of these effects.

We have demonstrated the algorithm on SPIM data, as this system was specifically designed to generate high-quality 3D multi-view fluorescence microscopy images of biological specimens. Since the technique provides optically sectioned fluorescence images with high signal-to-noise ratios, they are ideal for deconvolution-based image fusion. However, the image processing technique described here is not intrinsically limited to SPIM: in principle it could be applied to traditional wide-field, confocal, or two-photon data stacks, or even non-fluorescent or non-optical 3D imaging techniques. The requirements are 1) that the imaging system be able to generate 3D data stacks along multiple viewing directions, 2) that these views are not significantly distorted and have sufficient overlap, so that the pre-processing plus translation allows them to be registered, 3) that the imaging process be spatially invariant and quantitative (the processing would not be expected to work optimally with data obtained by, e.g. differential interference contrast (DIC), in which the response is a non-linear function of the sample properties), and 4) that the PSF can be determined. The last restriction could be relaxed if one were to substitute a “blind multi-view deconvolution” for the MVD-Wiener and/or MVD-MAPGG routines.

The fusion of multi-view SPIM images permits novel experiments in fields such as embryological gene expression. For instance, the Medaka embryo data shown in Fig. 6 was generated in a feasibility test for a large-scale screen to determine gene expression localization. It is evident from the results presented here that SPIM fusions can provide high-resolution information about 3D gene expression patterns throughout entire embryos.

Acknowledgments

The authors thank Christina Cardoso and Harald Saumweber for kindly donating the nuclei-labeled *Drosophila* sample, and Mirana Ramialison and Jochen Wittbrodt for the raw Medaka data.