

Dissertation
submitted to the Combined faculties
for the natural Sciences and for Mathematics
of the Ruperto-Carola University of Heidelberg, Germany
for the degree of Doctor of Natural Sciences

Put forward by
Dipl.-Phys. **Uroš Kržič**
Born in Slovenj Gradec, Slovenia
Oral examination: 8th July 2009

**Multiple-view microscopy
with light-sheet based fluorescence microscope**

Referees: Prof. Dr. Jörg Schmiedmayer, Vienna University of Technology
Prof. Dr. Bernd Jähne, Heidelberg University

Zusammenfassung:

Die axiale Auflösung jedes auf einem einzelnen Objektiv basierenden Lichtmikroskops ist schlechter als seine laterale Auflösung. Daher ist auch die Auflösung in einem konfokalen oder einem zweiphotonenabsorbierenden Fluoreszenzmikroskop entlang der optischen Achse schlechter als in der Fokalebene. Das Verhältnis der Auflösungen ist 3 bis 4 bei hohen numerischen Aperturen (NA 1,2 – 1,4) und kann für niedrige numerische Aperturen (NA < 0,2) sogar Werte von 10 bis 15 erreichen. Damit ist der Einsatz der konventionellen Lichtmikroskopie gerade für große Objekte eventuell sehr stark eingeschränkt. Die schlechte axiale Auflösung ist nicht ausreichend, um kleine Objekte innerhalb einer Zelle zu lokalisieren. Dazu kommt, dass große Objekte nicht komplett erfasst werden können. Die Beobachtung entlang mehrerer Raumrichtungen stellt sich dieser Aufgabe, indem sie Bildstapel desselben Objekts entlang verschiedener Winkel aufzeichnet. Diese unabhängig voneinander aufgezeichneten Bildstapel werden in einem nachfolgenden Prozess zu einem neuen Datensatz zusammengefasst.

Die Datenaufzeichnung entlang unterschiedlicher Raumrichtungen wurde am EMBL im Rahmen der Fluoreszenz-Lichtscheibenmikroskopie entwickelt (LSFM). Das LSFM ist bislang das einzige bekannte Mikroskop, an dem ein solches Konzept zur Datenfusion erfolgreich demonstriert werden konnte. In dieser Arbeit werden die Aspekte eines LSFM, die für die Aufnahmen entlang unterschiedlicher Raumrichtungen wichtig sind, charakterisiert. Außerdem wird die Implementierung eines LSFM ausführlich beschrieben. Wesentliche Aspekte werden sorgfältig diskutiert und in den allgemeinen Kontext bereits publizierter Arbeiten gestellt. Die Bilderfassung unterschiedlicher Objekte (u.a. Medaka Fisch, Fruchtfliege, Bäckerhefe) illustriert zwar die Grenzen aber vor allem die Möglichkeiten.

Abstract:

The axial resolution of any standard single-lens light microscope is lower than its lateral resolution. The ratio is approximately 3-4 when high numerical aperture objective lenses are used (NA 1.2 -1.4) and more than 10 with low numerical apertures (NA 0.2 and below). In biological imaging, the axial resolution is normally insufficient to resolve subcellular phenomena. Furthermore, parts of the images of opaque specimens are often highly degraded or obscured. Multiple-view fluorescence microscopy overcomes both problems simultaneously by recording multiple images of the same specimen along different directions. The images are digitally fused into a single high-quality image.

Multiple-view imaging was developed as an extension to the light-sheet based fluorescence microscope (LSFM), a novel technique that seems to be better suited for multiple-view imaging than any other fluorescence microscopy method to date. In this contribution, the LSFM properties, which are important for multiple-view imaging, are characterized and the implementation of LSFM based multiple-view microscopy is described. The important aspects of multiple-view image alignment and fusion are discussed, the published algorithms are reviewed and original solutions are proposed. The advantages and limitations of multiple-view imaging with LSFM are demonstrated using a number of specimens, which range in size from a single yeast cell to an adult fruit fly and to Medaka fish.

TABLE OF CONTENTS

1	Introduction.....	5
1.1	Optical microscopy.....	6
1.2	Trends in biological imaging.....	7
1.3	Fluorescence microscopy.....	7
1.4	Fluorophores.....	8
1.5	Fluorescent dyes.....	11
1.5.1	Fluorescent proteins.....	12
1.5.2	Quantum dots.....	13
1.6	Common fluorescence microscopy techniques.....	13
1.6.1	Wide-field fluorescence microscope.....	14
1.6.2	Confocal microscope.....	19
1.6.3	Two-photon microscope.....	22
1.6.4	Other optical sectioning microscopes.....	24
1.6.5	Lateral vs. axial resolution.....	25
1.6.6	Super-resolution methods.....	27
2	Light-sheet based fluorescence microscope (LSFM).....	31
2.1	Use of light-sheets in light microscopy.....	33
2.2	Basic principles.....	35
2.2.1	Detection unit.....	35
2.2.2	Illumination unit.....	39
2.2.3	Single plane illumination microscope (SPIM).....	41
2.2.4	Digital scanned laser light sheet microscope (DSLIM).....	51
2.2.5	Specimen translation and rotation.....	52
2.3	Specimen preparation and mounting.....	53

2.4	LSFM application: imaging of hemocyte migration in a <i>Drosophila melanogaster</i> embryo.....	57
2.4.1	Drosophila m. hemocytes	57
2.4.2	Drosophila transgenes	59
2.4.3	Automated hemocyte tracking	60
2.4.4	Laser induced wounding and wound induced hemocyte migration	62
3	Multiple-view microscopy with LSFM.....	65
3.1	Motivation.....	66
3.2	Multiple-view imaging in microscopy	69
3.3	Multiple-view microscopy with LSFM	71
3.4	Multiple-view image acquisition.....	72
3.5	Multiple-views image fusion.....	76
3.5.1	Image formation and sampling	78
3.5.2	Preprocessing.....	79
3.5.3	Image registration.....	80
3.5.4	Final image fusion	96
3.6	Examples of multiple-view microscopy on biological specimens.....	124
4	Summary and outlook.....	133
	Bibliography	135
	Abbreviations.....	145
	Table of figures	147
	Acknowledgements.....	151

1 INTRODUCTION



The human mind prefers something, which it can recognize to something for which it has no name, and, whereas thousands of persons carry field glasses to bring horses, ships, or steeples close to them, only a few carry even the simplest pocket microscope. Yet a small microscope will reveal wonders a thousand times more thrilling than anything, which Alice saw behind the looking-glass.

David Fairchild, American botanist
The World Was My Garden (1938)

Sight is regarded as the single most important channel through which the human mind perceives its surrounding world. It seems to be such a vital source of information for a seeing man, that objects are often not considered existent unless they can be visualized: seeing is believing. Optics (meaning “look” in ancient Greek) and simple optical apparatuses seem to be the oldest tools that allowed the humans to perceive the world beyond the limits of a naked human eye. The oldest known lens¹ was unearthed in the region that is considered the cradle of the civilization and is believed to be more than 3000 years old. “Burning glasses” or “looking glasses” were often used in the Roman engravers’ workshops², while simple “flee glasses” were a common attraction on medieval fairs and excited general wonder and curiosity. By the 14th century, optics entered the common people’s lives through the wide spread of spectacles.

In the late 16th century, spectacle producers of the Low Countries recognized that more powerful optical apparatuses can be realized by a combination of multiple lenses. These instruments (described by Robert Hook as “artificial organs” improving our natural senses) were turned against the sky, the world around us and against our own bodies, vastly improving our understanding the life and the universe around us.

Obviously, optics has advanced a great deal since then. A number of instruments were introduced that in fact do not share anything with the crude devices of Jansen and Leeuwenhoek, other than

¹ Named Nimrud lens after an ancient Assyrian city near the location of modern Mosul in Iraq, where it was found. The analysis of the lens by D. Brewster was published in *Die Fortschritte der Physik* in 1852.

² Most famous are writings by Seneca and Pliny, describing the lens used by an engraver in the ancient Pompeii.

the name *microscope*. However, somewhat opposing E. Abbe's visions³, optical microscope stayed the most popular tool in the life sciences throughout the previous century. Nevertheless, light microscopy was "reinvented" in the recent three decades by the advent of the fluorescence microscopy and the techniques enabled by fluorescence.

One of such novel fluorescence microscopes, light-sheet based fluorescence microscope (LSFM), is in the focus of this work. Aspects of LSFM, important for its construction and understanding of its operation, are described in Chapter 2, together with a selection of LSFM's applications to biological problems. Chapter 3 concentrates on multiple-view LSFM imaging, i.e. imaging the same specimen along multiple axes and the subsequent digital fusion of the images into a single image, in order to improve both, resolution and image completeness. But first, elements of microscopy important for understanding of chapters 2 and 3 are briefly recapitulated in this section.

1.1 Optical microscopy

Optical microscopy operates mainly with the visible light, which spans the spectral range between 350nm and 800nm. However, it is not uncommon to use wavelengths as low as 300nm and as high as 1100nm. The sources of light range from light bulbs to LEDs and lasers. The paths of the light beams are easily controlled using mirrors, apertures and lenses. Finally, light is commonly recorded with photomultipliers as well as CCD cameras. In fact nowadays, the basic technology for optical microscopy is very mature, readily available and usually of excellent quality.

Optical microscopy is most commonly associated with transmitted or reflected light. The contrast is generated by absorbers or scatterers in the specimen or by taking advantage of a specimen's birefringent properties. However, modern optical microscopy takes advantage of all properties of light (wavelength, polarization, momentum) and discriminates against essentially all manners, in which a specimen influences these properties. Hence a wealth of different contrasts is available that allows one to probe many properties of a specimen.

In life sciences related optical microscopy, the most important property of a specimen is fluorescence. This is the process of absorbing at least one photon with a well-defined energy and emitting a photon with a different energy within a statistically determined time frame. Fluorophores can be attached to various biological compounds (e.g. lipids, proteins, nucleic acids) and provide a specific labeling; i.e. only compounds of a certain class contribute to an image. The specificity allows one to compare the spatial and temporal distribution of different targets and to relate them to biological processes. Fluorescence may provide a relatively low signal (only 0.0001% of the photons focused on the specimen result in a fluorescence photon) but the contrast is very high.

The main advantage of optical microscopy over electron microscopy is that it allows researchers to observe live specimens. Other advantages are the simple specimen preparation, the cheap instrumentation and the relatively easy access to various types of equipment. The main

³Ernst Abbe in 1878 farsightedly noted: "Perhaps in some future time the human mind may succeed in finding processes and in conquering natural forces, which will open important new ways to overcome in an unforeseeable manner the limitations which now seem insurmountable to us. This is my firm belief. But I think that the tools which some day will aid mankind in exploring the last elements of matter much more effectively than the microscope, as we now know it, will probably have no more in common with it than its name." [172]

disadvantage of optical microscopes compared to electron microscopes is their lower resolution. The resolution is essentially limited by the properties of the optical system and will be close to the wavelength of the excitation light.

More recent developments such as e.g. confocal fluorescence microscopy and PALM or STORM sample a specimen. The important issue becomes the precision with which the source of the fluorescence emission is localized. The sampled fluorescence intensity as a function of the position at which it was recorded is used to assemble an image in a computer. Sampling methods are always slower than modes that record images and hence can use cameras to record millions of picture elements in parallel. Sampling usually also results in very high radiation doses to which a specimen is exposed.

One of the most important properties of optical systems is their optical sectioning capability. Only those methods that rely on two independent processes of exciting a fluorophore and collecting the emitted light provide the optical sectioning capability. Conventional fluorescence microscopy does not provide it, while confocal (theta) and multi-photon-absorption fluorescence microscopy do.

1.2 Trends in biological imaging

The main challenge in modern biology is to observe physiologically relevant, live specimens with a high spatial resolution, a high temporal resolution, a high specificity and multiple times over extended periods of time. In addition, it is absolutely essential to carefully assess the impact of an experiment on the physiology of the specimen. The observation as well as the optical manipulation of extended biological specimens suffers from at least two severe problems. 1) The specimens are optically dense, i.e. they scatter and absorb light. Thus, the delivery of the probing light and the collection of the signal light tend to become inefficient. 2) Many biochemical compounds apart from fluorophores also absorb light and suffer degradation of some sort (photo-toxicity), which induces malfunction or death of a specimen [1]. The situation is particularly dramatic in conventional and confocal fluorescence microscopy. Even though only a single plane is observed, the entire specimen is illuminated. Recording stacks of images along the optical z-axis thus illuminates the entire specimen once for each plane. Hence cells are illuminated 10-20 [2] and fish embryos even 100-300 times more often than they are observed.

Moreover, most modern optical technologies (microscopy, optical tweezers [3], laser nanoscalpel [4]) are applied to two-dimensional cellular systems, i.e. they are used in a cellular context that is defined by hard and flat surfaces. However, physiological meaningful information relies on the morphology, the mechanical properties, the media flux and the biochemistry of a cell's context found in live tissue [1]. A physiological context is certainly not found in single cells cultivated on cover slips. It requires the complex three-dimensional relationship of cells cultivated e.g. in an ECM-based gel, on collagen or in naturally developing small embryos of flies or embryos and, of course, in tissue sections [1].

1.3 Fluorescence microscopy

Fluorescence microscopy [e.g.5] is based on fluorophores [6], a special group of chemicals that fluoresce, i.e. they emit a photon within nanoseconds after they absorbed another photon typically with a shorter wavelength. The difference between the wavelength of absorbed and emitted

photons, the Stokes shift (see chapter 1.4), is essential for fluorescence imaging. It allows one to discriminate the illumination light while transmitting most of the fluorescence using an appropriate spectral filter. Once scattered and reflected light is filtered the fluorophores remain visible and are ideally seen in a standard fluorescence microscope as bright sources of light on a dark background. Contrast is superior to absorption microscopy techniques, where structures of interest are seen as dark areas on bright background.

Fluorescent microscopy is especially potent in combination with specific staining [7]. Only a small fraction of biological molecules fluoresce naturally (autofluorescence). Fluorophores used for microscopy are typically introduced into a biological specimen from outside (staining) or encoded into an animal's genome so that they specifically associate with selected proteins, cells or tissues. The expression and localization of the selected proteins can thus be observed, and consequently the movement and the proliferation of the selected cells can be followed and the development of chosen tissues can be visualized. Thanks to highly specific absorption and emission spectra of different fluorophores, more than one fluorophore can be used in a single specimen, simultaneously highlighting different proteins or tissues. The specimen's autochthonous fluorescence normally adds unwanted signal and it is often regarded as a nuisance. However, autofluorescence can also be useful when autofluorescing structures are being studied.

High contrast and sensitivity combined with specific staining made fluorescence based microscopy a mainstay of modern biological imaging. Like the rest of the optical microscopy techniques, fluorescence microscopy is applicable to living cells, tissues and animals, allowing the observation of dynamic processes where they happen naturally.

1.4 Fluorophores

A fluorophore [6,8] (sometimes also fluorochrome) is a molecule or a part of a molecule, which exhibits fluorescence. Typically they are compounds with some degree of conjugated double bonds or aromatic rings with π bonds that distribute outer orbital electrons over a wider volume. As a rule of thumb, more conjugated bonds in a fluorophore mean lower fluorescence excitation and emission wavelength peaks (redder spectra) and better efficiency (quantum yield).

When a fluorophore absorbs a photon, it takes up its total energy. The energy of a photon (E) is proportional to its frequency (ν) and inversely proportional to its wavelength in vacuum (λ_0):

$$E = h \cdot \nu = \frac{h \cdot c_0}{\lambda_0} \quad (1.1)$$

Where h is Planck constant and c_0 is speed of light in vacuum.

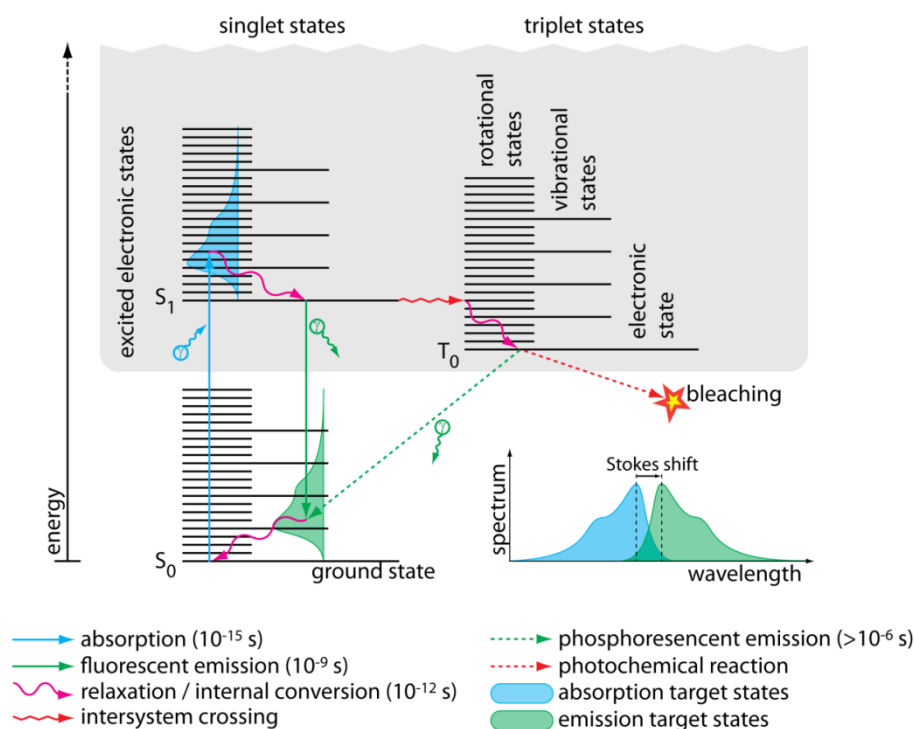


Figure 1: Jablonski diagram of a typical fluorophore. The diagram shows the different energy levels of a fluorophore and the most common transitions between them. Each energy level has a distinct electronic, vibrational and rotational state. Dense rotational and vibrational states superimpose the ground electronic state (S_0) to form a ground energy band and the states superimpose the excited electronic states (singlet S_1 and triplet T_0) to form the excited energy band. The forbidden transitions between singlet and triplet states are much less likely than transitions between exclusively singlet or triplet states. The blue line represents the photon absorption. The other arrows denote the main subsequent energy relaxation pathways that end either in the ground state or in a photo-bleaching event. In the bottom-right corner is an example for a typical absorption (blue) and a typical emission (green) spectrum. The time durations in the brackets illustrate the typical time required for a transition when all conditions are met. (Follows [5])

Upon absorption, the fluorophore undergoes a transition into a higher energy, excited quantum state. The energy of the photon is stored by the higher vibrational and rotational states and, if sufficiently large, it can also lift an electron from one of the outer filled molecular orbitals into an higher energy unoccupied electronic orbital. The latter process occurs on a scale of femtoseconds (see Figure 1).

Once the fluorophore is excited, it can shed the additional energy in a number of different ways. The energy distribution of rotational and vibrational states is very dense, forming an energy band of states that are connected with quick and highly probable transitions. The probable path of energy relaxation from any state in a band is thus rotational and vibrational relaxation, where the molecule transfers the superfluous energy to the neighboring molecules (in biological specimens usually water molecules), while “drifting” to the bottom of the band. At the same time, the molecule undergoes internal conversions, an isoenergetic transformation of a molecule’s configuration from high-electronic and low-vibrational state to a lower-electronic and higher-vibrational state with a similar energy. Rotational and vibrational relaxation and internal conversion are very brief processes and a fluorophore typically relaxes to thermic vibrational and rotational states within picoseconds.

An essential property of all fluorophores used for imaging is that the energy bands of the electronic ground state and electronic excited states are separated by an energy gap that is wide enough to make non-radiative relaxation highly improbable. The excited fluorophore gets “stuck” at the bottom of the excited energy band until it relaxes by expulsion of a photon, i.e. fluorescence emission. The typical average life-time of an excited fluorophore before it undergoes a fluorescent emission is in the range of several nanoseconds, which makes it much less likely than a direct relaxation via internal conversion, if the latter is possible. Only molecules with well separated energy bands therefore fluoresce.

The emitted photon is only weakly related to the absorbed photon. Due to the fluorophore’s relaxation between the absorption and emission, the emitted photon’s wavelength does not depend on the wavelength of the absorbed photon. It is defined primarily by the energy gap. Likewise, the direction of the emitted photon is not related to the direction of the absorbed photon⁴.

Since a fraction of total absorbed energy is released non-radiatively, the emission spectra are typically shifted to longer wavelengths. The difference between absorption and emission peaks is called the Stokes shift and typically amounts to a fraction of an electron volt (0.1 – 0.5 eV).

Absorption spectra can be explained by the energy differences between ground state and different states in the excited energy band. Due to a quick energy relaxation inside an energy band, fluorophores normally only absorb and emit a photon when they are at the bottom of the band, i.e. near the energy of a purely electronic state. The probability of a fluorophore being excited into any of the vibrational/rotational states in the excited band is roughly equivalent (if a photon with the corresponding energy is present). The probability of the fluorophore being excited into a state

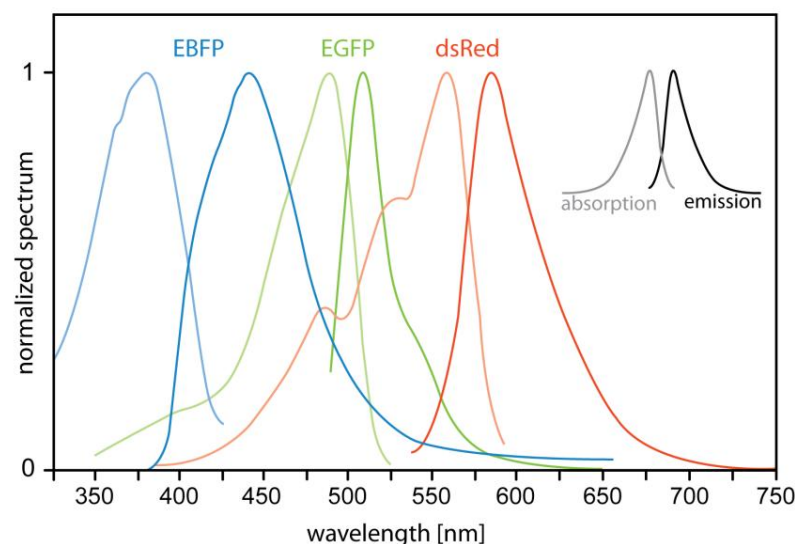


Figure 2: Examples of fluorescent protein spectra. Absorption (lighter lines) and emission (darker lines) spectra of three different proteins are shown: EBFP – enhanced blue fluorescent protein (blue), EGFP – enhanced green fluorescent protein (green) and dsRed. While the first two proteins were created by mutating the wild type GFP (green fluorescent protein) isolated from *Aequorea Victoria* jellyfish [2], dsRed was discovered in Anthozoan genus *Discosoma* [3]. Spectra are based on [4].

⁴The polarization of the emitted photon in the reference system of the fluorophore is preserved. The degree of total polarization dispersion therefore depends on the angular mobility of the fluorophores in the specimen and is thus a source of information [173].

with the energy between E and $E + dE$ is therefore roughly proportional to the density of the states in that energy interval. Similar reasoning applies also to fluorescence emission, when the fluorophore relaxes from the bottom of the excited band to one of the vibrational/rotational states in the ground band. Since ground and excited energy bands are formed by the same vibrational and rotational states superimposed on top of two different electronic levels, the density of states inside both bands have very similar profiles. This is why the absorption and emission spectra of a typical fluorophore look mirrored (see inset of Figure 1 and Figure 2).

The energy release pathway described above is the preferred, but not the only one. An electron in the excited, higher energy orbital can undergo an intersystem crossing, an unlikely forbidden transition that effectively inverts its spin. A fluorophore enters a triplet state and it can only drop back to a singlet state by an inverse, i.e. a forbidden transition. A triplet energy band overlaps with the excited singlet energy band, but it is shifted to slightly lower energies (Figure 1). After a fluorophore enters a triplet state, it quickly relaxes its vibrational energy which makes its reentry into the excited singlet state highly improbable. The fluorophore stays trapped in the triplet state until it undergoes another forbidden transition with the simultaneous emission of a photon, i.e. phosphorescence. An average fluorophore stays locked in its triplet state for more than a microsecond, before it eventually relaxes. During this time it can absorb another photon, which further delays its release to a ground singlet level. Such a molecule is thus temporarily removed from a pool of quickly cycling fluorophores and will miss a couple of thousand absorption/emission cycles before it finally returns to a singlet state. This is especially disturbing in confocal and two-photon microscopy, where only a couple of microseconds are dedicated to the fluorescence measurement in every volume element. A fluorophore entering a triplet state can thus be considered practically lost for detection.

Last but not least, the high energy stored by excited fluorophores trapped in a triplet state allows them to react chemically with molecules in their vicinity. Such photochemical reactions disturb the physiology of living cells (induced photo-toxicity) and permanently damage the fluorophores, permanently removing them from the pool of fluorophores in the specimen. The process is called photo-bleaching and is one of the most fundamental challenges of the modern fluorescence microscopy.

1.5 Fluorescent dyes

Synthetic fluorescent dyes are systematically used in microscopy since the late 19th century. By that time the dyes were used to increase the contrast of histological sections by selectively coloring structures of interest. These dyes absorbed selected bands of visible light, and looked colorful under a microscope. Fluorescence emission excited less interest, nevertheless several synthetic dye classes synthesized during this period, based on the xanthene and acridine heterocyclic ring systems, proved to be highly fluorescent and provided a foundation for the development of modern synthetic fluorescent probes. Most notable among these early fluorescent dyes were the substituted xanthenes, fluorescein and rhodamine B, and the biaminated acridine derivative, acridine orange.

New families of fluorophores with higher quantum yields for more intense fluorescence have been developed in recent decades, e.g. the cyanine dyes (Cy2, Cy3, Cy5...) and Alexa Fluor by Invitrogen (Life Technologies Corporation, Carlsbad, USA). They offer a broad range of

fluorescence excitation and emission spectra, ranging from the ultraviolet and deep blue to the near-infrared regions. They are routinely conjugated to phalloidin, lectin, dextran, streptavidin, avidin, biocytin, and a wide variety of secondary antibodies.

A notable milestone of specific fluorescence staining was a technique for fusing fluorescent dyes with mammalian antibodies [9], devised in 1940 by Albert Coons. Antibodies (or immunoglobulins) are proteins found in the blood of vertebrates and are involved in their humoral immune system i.e. the neutralization of unidentified proteins. When a vertebrate (usually a mammal, e.g. mouse, rat or rabbit) is injected with a protein of interest, its immune system produces antibodies that bind to the introduced protein (i.e. antigen). Once the antibodies are isolated from the animal's blood and fused with fluorescent dyes, they are used to specifically stain the protein of interest. In the years following the pioneering work of Albert Coons, immunostaining was developed into one of the most convenient and widely used methods for specific fluorescent marking. Its importance was acknowledged in 1984 when Georges Köhler, César Milstein, and Niels Kaj Jerne were awarded a Nobel Prize in Medicine for their contribution to the development of monoclonal antibodies.

A recent development in the field of synthetic fluorophores is the development of dyes that reveal intercellular ion concentrations (e.g. calcium [10]). These probes bind to a target ion that produces change in their spectral properties and are thus referred to as *spectrally sensitive indicators*. Finally, cell permeant fluorophores targeting specific intracellular organelles have been engineered, such as the mitochondria (MitoTracker, Invitrogen), lysosomes (LysoTracker, Invitrogen), Golgi apparatus, and endoplasmic reticulum (DiOC(6)) [8,11].

1.5.1 Fluorescent proteins

Over the past decade, the isolation of naturally occurring fluorescent proteins and the development of mutated derivatives made fluorescent proteins a widely used fluorescent marker [12]. The first fluorescent protein to be purified, sequenced and cloned was the green fluorescent protein (GFP), isolated from the North Atlantic jellyfish *Aequorea victoria* [13-15]. A large number of variants was produced by the mutagenesis of GFP since then, exhibiting improved folding (EGFP) and modified absorption and emission spectra (BFP, CFP, EYFP, RFP). Proteins fluorescing in the red part of the visible spectrum, unattainable by mutagenesis of GFP, were isolated from other marine species, (DsRed from sea anemone *Discosoma striata* [16], HcRed from purple anemone *Heteractis crispa*).

The Nobel Prize for chemistry in 2008 was awarded to Osamu Shimomura, Martin Chalfie and Roger Y. Tsien “for discovery and development of the green fluorescent protein, GFP.” [17]

A strong advantage of fluorescent proteins over synthetic fluorescent dyes is that they are synthesized directly by the animal or plant once a fluorescent protein's gene has been correctly introduced into an organism's genome. Furthermore, a fluorescent domain is often attached to a protein. The protein's location and level of expression are then visualized by fluorescence microscopy.

Newly discovered photoactivatable fluorescent proteins become fluorescent only after they are activated by absorbing light with a specific wavelength [18]. This wavelength is different (typically shorter) than the one used for subsequent fluorescence excitation and causes structural

changes of the fluorophore. This, in turn, results in an increased fluorescence or a shift in fluorophore's absorption or emission spectra (photoconversion). Most widely used photoactivatable fluorescent proteins include photoactivatable green fluorescent protein (PA-GFP) [19], Kaede [20], Dronpa [21] and EosFP [22]. These probes proved useful in fluorescence studies involving selective activation of fluorescence in specific target regions and the subsequent kinetic analysis of diffusion mobility and compartmental residency time of fusion proteins [23]. Furthermore, they have recently been used to push the boundaries of far field fluorescence localization well over Abbe's diffraction limit (section 1.6.6).

1.5.2 Quantum dots

Several nanometers sized crystals of purified semiconductors known as quantum dots (QDs) are emerging as a potentially useful fluorescent labeling agent for living and fixed cells in fluorescence microscopy [7,24,25]. These probes have significant benefits over organic dyes and fluorescent proteins, including long-term photostability (e.g. they were observed for four months in lymph nodes of mice [26]), high fluorescence intensity, and configurable emission spectra at a single excitation wavelength. The most widely used semiconductor for quantum dots used in biological imaging is cadmium selenide (CdSe). The emission spectrum is defined by the energy gap between the valence and conduction band of a bulk semiconductor, however due to the quantum confinement, the peak wavelength of a QD can be tuned by the crystal's size. As a general rule, smaller crystals will emit light that is more blue-shifted from the band gap energy of the bulk semiconductor. QD is covered by several layers of coating that improve its optical properties, make it more biologically inert and functionalize its surface to induce selective binding. The actual disadvantage of the quantum dots is their tricky delivery into a cell or organism. It is so far limited to injection or ingestion by endocytotic cells (macrophages, lymphocytes). For an extensive review on the use of quantum dots in biological imaging see [25].

1.6 Common fluorescence microscopy techniques

Almost a century ago, scientists at Carl Zeiss company in Jena were experimenting with use of shorter wavelengths to increase the resolution of transmission light microscopes. When irradiated with near ultra-violet light, which is invisible for human eye, parts of some biological specimens were reported to glow in green colors (the discovery is attributed to August Köhler, inventor of transmission illumination technique that bears his name). This was due to the Stokes shift (section 1.4) by the fluorescent compounds in the specimen, which shifted invisible UV light into the visible spectrum. The first epi-fluorescence microscope (see section 1.6.1) was reported in 1925 by A. Policard and A. Paillot who used it to study different biological processes involving autofluorescence. However, fluorescence microscopy didn't trigger broader interest until Coons' invention of immunostaining in 1940.

Fluorescence microscopy witnessed a revival since the early 1970s. This was not due to one single invention, but rather thanks to progress in many different fields, ranging from optics (confocal microscopy in 1957, laser in 1959, fluorescence confocal microscopy in 1984 [27]), detectors (charge-coupled device – CDD in 1969), mechanical and electrical engineering, computing and finally modern genetics and fluorescent proteins (green fluorescent protein in 1994). The combination of these advances has revolutionized modern biological imaging and

powered a rapid development of new microscopic techniques, the most important of which will be described in the following sections.

The ultimate task of fluorescence imaging is to determine a density distribution of the fluorophores in a fluorescent specimen. Knowing how fluorophores relate to the protein or other structures of interest, the location and density of the fluorophores can tell a researcher a great deal about the proteins and the processes they are involved in. The basic principles of fluorescence measurement are the same for all fluorescence microscopes: fluorophores are first excited and the intensity of the emitted fluorescence is measured. However, each of the microscopy techniques applies different mechanisms to achieve spatial discrimination at resolution close to Abbe's limit. Some of the emerging "super-resolution" microscopy methods are summarized in section 1.6.6.

1.6.1 Wide-field fluorescence microscope

The simplest fluorescence microscope is a standard wide-field microscope, equipped with spectral filters that shape the illumination and detection spectra. The emitted fluorescence is collected by an objective lens, which combined with a tube lens creates an image on an image sensor (e.g. a CCD camera), or, through an ocular and the eye's lens, on an eye's retina.

Since there is no correlation between the direction of the photon that is absorbed by a fluorophore and the one that is eventually emitted, the fluorescence is in general radiated isotropically. It can therefore be detected by the same set of optics that was used for illumination (unlike in standard transmission microscope). The design where the same objective lens is used for both, illumination and detection of fluorescence, is called epifluorescence microscope (Figure 3).

Shaping of the illumination and detection spectra in an epifluorescence microscope is done by the *filter-cube*, which consists of two spectral filters and a dichroic mirror. The transition wavelength of the latter is chosen such that it reflects most of the excitation wavelengths while transmitting most of the fluorescence. The excitation and detection filters additionally optimize the excitation and detection spectra to minimize the illumination intensity at the wavelengths that do not excite the fluorophores and reduce the detection of light other than fluorescence.

1.6.1.1 Resolution of a standard fluorescence microscope and its point spread function

The resolving power of an optical instrument is fundamentally limited by the wave nature of the light. In microscopy, resolving power is commonly described in terms of optical resolution, i.e. the shortest distance between two point-objects that can still be resolved as separate entities by the microscope. Resolution of standard microscopes is usually in the range of the wavelength of the light used.

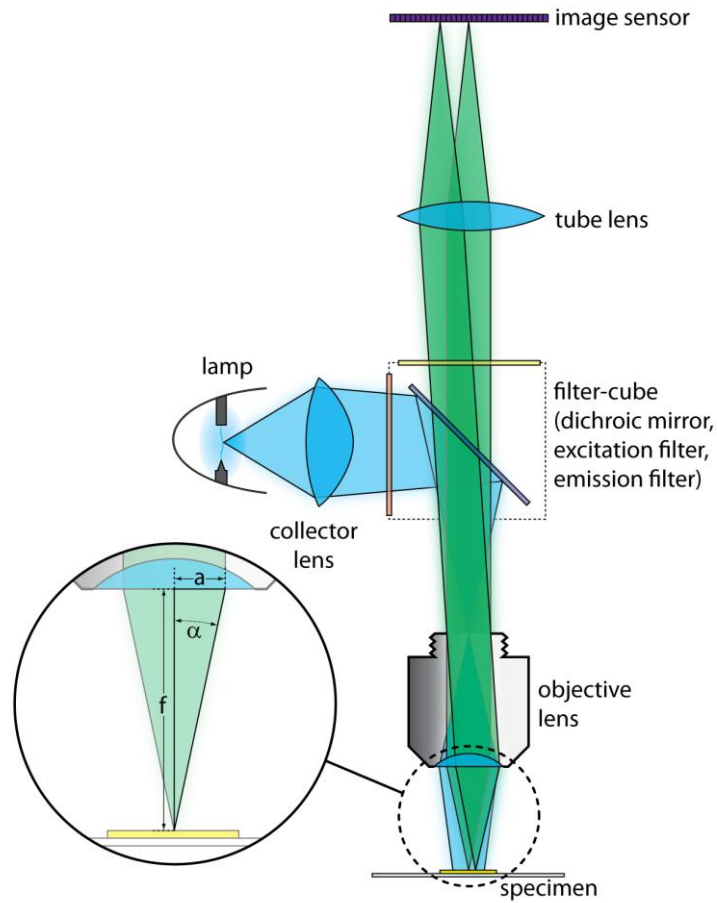


Figure 3: Diagram of an epifluorescence microscope. Essential parts of an epifluorescence microscope are shown with illumination (blue) and detection (green) pathways.

There are a few different exact definitions of optical resolution and ways to calculate it. A novel simple and generally applicable method by Stelzer and Grill, similar to Heisenberg's uncertainty inequality (therefore called Stelzer-Grill-Heisenberg or SGH theory), yields the following expressions for the optical resolution of a microscope [28,29]⁵ as defined by the Rayleigh criterion [30]:

$$\sigma_{xy} = \frac{\lambda}{\sqrt{3 - 2 \cos \alpha - \cos 2\alpha}} \quad (1.2)$$

$$\sigma_z = \frac{\lambda}{1 - \cos \alpha} \quad (1.3)$$

where σ_{xy} and σ_z are resolutions in the focal plane (lateral resolution) and along the optical axis (axial resolution), respectively, λ is light wavelength and α is the collecting half angle of the microscope's objective lens (Figure 3, inset). A common way of describing the objective lens' aperture is the numerical aperture: $NA = n \sin \alpha$, where n stands for refractive index of the immersion medium and $\sin \alpha$ is the angular aperture.

⁵ The original formulas published in the cited references have been multiplied by a coefficient $2\pi/\sqrt{3}$ to make the results consistent with the definition of optical resolution based on the Rayleigh criterion.

Graphs of both resolutions as functions of numerical aperture can be seen in the Figure 4. Note that the axial resolution is always worse than the lateral resolution. They become equal only in case of $\alpha = \pi$, which would correspond to an *isotropic microscope* coherently collecting light from all around the specimen (in this case “axial” and “lateral” lose their meaning). Such a device currently seems hardly achievable. When the numerical aperture decreases towards zero, the axial resolution diverges faster (second degree divergence) than the lateral resolution (first degree divergence).

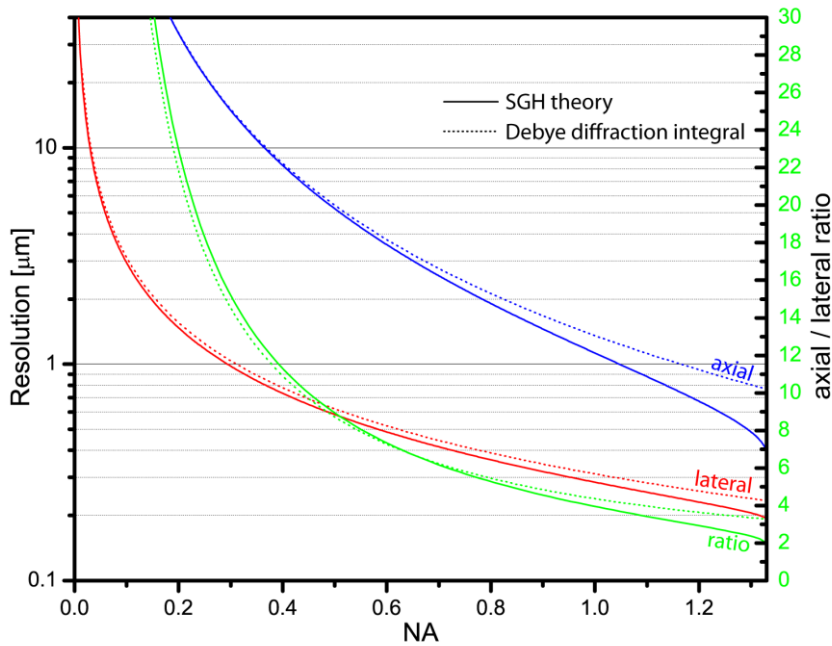


Figure 4: Resolution of a wide-field microscope. The lateral (red line) and axial (blue line) resolutions of a standard wide-field microscope are shown together with their ratio (green line). The resolutions were calculated using standard diffraction integral (often labeled Born-Wolf approximation, referring to [30]; dotted line) and Stelzer-Grill-Heisenberg theory [28] (solid line), which gives more trustworthy results at high NAs. The calculations were done for water as immersion medium (refraction index $n = 1.333$) and light wavelength corresponding to GFP emission peak ($\lambda_0 = 510nm$).

In a linear and space invariant image formation model (for more, see Chapter 3.5.4), the imaging properties of a microscope are most generally described in terms of a point spread function (PSF). The PSF is a response of an optical setup to a point-like source of light. The *amplitude* PSF is in general a complex number $h(x, y, z) = |h(x, y, z)|e^{i\varphi(x, y, z)}$, where the phase part relates the phases of the electric field oscillation in the source object and in the image. Our eyes, photographic film and electronic image sensors only detect the intensity of the light. The image collected by the image sensor is therefore determined by the *intensity PSF*:

$$H(x, y, z) = h^*(x, y, z) \cdot h(x, y, z) = |h(x, y, z)|^2 \quad (1.4)$$

The intensity PSF is normalized: $\iiint H(x, y, z) dx dy dz = 1$. From a quantum-electrodynamics point of view, the intensity PSF $H(x, y, z)$ refers to a probability density that a photon emitted from the origin of the object space is detected by a point-detector at a position that corresponds to coordinate (x, y, z) . Sometimes, an inverse understanding is used and the PSF is interpreted as a

probability that a photon detected at the origin of the image space originated from a coordinate $(-x, -y, -z)$ in the object space.

We can regard an extended light source as a distribution of point sources, each of them creating a spot with a shape of intensity PSF in the image. As fluorescent light is intrinsically incoherent, intensity spots produced by individual fluorophores add up to the final intensity image. Such an image is therefore a three-dimensional convolution of the fluorophore distribution in the object space $O(x, y, z)$ and the imaging system's intensity PSF:

$$I(x, y, z) = O(x, y, z) * H(x, y, z) \quad (1.5)$$

Only light radiating from the specimen into a certain solid angle is collected by a microscope objective lens. The shape of the solid angle is described by the aperture, i.e. a fictitious window in a screen that blocks the rest of the light. In microscopy, the aperture has usually a round shape and its size is often defined by the size of the objective's front lens (Figure 3, inlet).

An approximate shape of the intensity PSF can be analytically expressed for some common aperture shapes using the Kirchhoff diffraction theory. If waves originating from an object are collected through a circular aperture with a radius a that lies a distance f away along the z axis, our detection system will express the following PSF [30]:

$$h(u, v) = A_0 e^{i\left(\frac{f}{a}\right)^2 u} \int_0^1 J_0(v\rho) e^{-i\frac{1}{2}u\rho^2} \rho d\rho \quad (1.6)$$

Where A_0 is a normalization factor, J_0 is a zero order Bessel function, while u and v are dimensionless variables defined as

$$u = \frac{2\pi n}{\lambda} \left(\frac{a}{f}\right)^2 z = \frac{2\pi NA^2}{\lambda n} z \quad \text{and} \quad v = \frac{2\pi n a}{\lambda f} r = \frac{2\pi NA}{\lambda} r \quad (1.7)$$

Here, NA stands for numerical aperture of the imaging system $NA = n \frac{a}{f}$, λ is wavelength of the detected light, n refractive index of the medium and $r = \sqrt{x^2 + y^2}$ is the distance from the optical axis. The derivation was based on assumptions of paraxial approximation, i.e. $f \gg a \gg \lambda$ and $a^2 = \lambda f$. It is expected to give reliable results only at low NAs.

The shape of the resulting intensity PSF, $H(r, z) = |h(r, z)|^2$ for three different values of NA can be seen in Figure 5. Note that the amplitude part of the solution of the integral (1.6) in dimensionless coordinates (1.7) does not depend on any additional parameters. Wavelength λ and NA influence only how the solution is mapped from dimensionless into real-space coordinates as defined by (1.7). As expected, image of the PSF expands linearly with decreasing NA in lateral dimension (x and y axes) and quadratically in axial dimension (z axis).

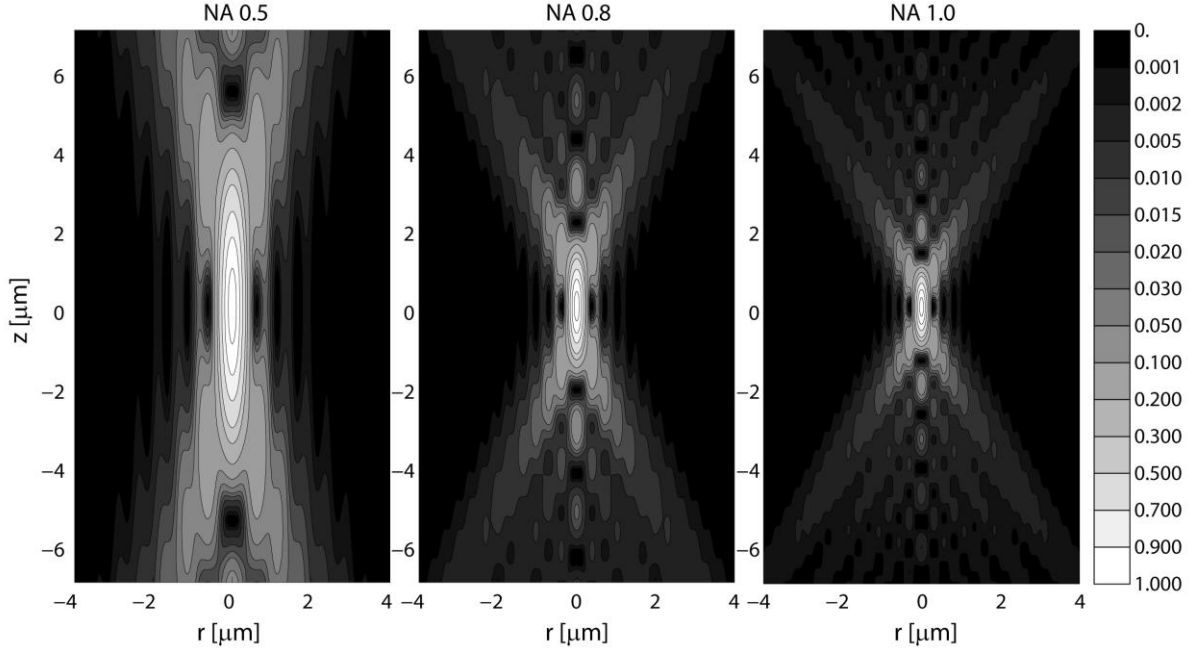


Figure 5: Intensity point spread function of a wide-field microscope. The PSF shapes for three different numerical apertures are shown. The calculation (section 1.6.1.1) is based on the Debye diffraction integral [30], water is the immersion medium (refractive index $n = 1.333$) and the wavelength of the light corresponds to the GFP emission peak ($\lambda_0 = 510nm$). The values are maximum-normalized. Note that all three shapes are similar, their extent decreases linearly with the NA along the r -direction and quadratically with the NA along the z -direction with increasing NA. The contour lines were chosen as in [30].

PSF Integral (1.2) can be analytically solved in a couple of interesting cases. The radial profile of the PSF in the focal plane of our imaging system is obtained by setting $u = 0$:

$$H^{wf}(0, v) = |A_0|^2 \left(\frac{2J_1(v)}{v} \right)^2 \quad (1.8)$$

where J_1 is the first order Bessel function. This is the formula of the Airy disc. The most commonly used definition of optical resolution, the Rayleigh criterion, defines the optical resolution as a distance between two point light sources where center peak in the image of the first source overlaps with the first minimum in the image of the second source. Considering that the first zero of the fraction in parenthesis in the equation (1.8) is at $v = 3.83$ and the second equation (1.7), the famous diffraction limit formula is derived:

$$\sigma'_{xy} = \frac{3.83}{2\pi} \frac{\lambda_0}{NA} = 0.61 \frac{\lambda_0}{NA} \quad (1.9)$$

Here λ_0 stands for vacuum light wavelength $\lambda_0 = n \lambda$.

Axial profile of the PSF is derived from (1.6) by setting $v = 0$:

$$H^{wf}(u, 0) = |A_0|^2 \left(\frac{\sin \frac{u}{4}}{\frac{u}{4}} \right)^2 \quad (1.10)$$

The first zero crossing of the sinc function in the brackets occurs at $u = 4\pi$ and the axial optical resolution as defined by Rayleigh criterion is therefore

$$\sigma'_z = 2 \frac{\lambda_0 n}{NA^2} \quad (1.11)$$

However, as mentioned above, these formulas are only expected to give reliable results at low NAs. Shapes of both functions are shown in the Figure 4 (dotted line) together with functions derived from SGH theory, which gives trustworthy solutions over the entire range of NA.

1.6.1.2 Contrast, depth of field and optical sectioning

Unlike in an idealized optical system, the resolving power of real imaging apparatuses is determined by the contrast⁶ they provide, which in turn is influenced by a number of parameters: sampling rate (image pixel size), dynamic range and signal-to-noise ratio [31]. In an epifluorescence microscope, the contrast is fundamentally limited by the light originating from outside the focal volume. An epifluorescence microscope does not discriminate against the out-of-focus light: a wide-field microscope's PSF (Figure 5) is not limited to the focal plane ($z = 0$) but rather stretches out to infinity. The integral of the PSF over any plane parallel to the focal is constant (see [32]):

$$E_{\Sigma}^{wf}(z) = \int_0^{\infty} H^{wf}(r, z) r dr = const. \quad (1.12)$$

A fluorophore anywhere below or above the focal plane therefore contributes the same integral intensity to the image; only if a fluorescent object is in or close to the focal plane it creates a distinguishable feature, otherwise it produces a blurred artifact and/or increases the homogeneous image background.

The background intensity thus depends on a fluorescent specimen's total thickness. While an epifluorescence microscope generates relatively high contrast images of thin specimens (e.g. single cells on a glass slide), the features in focus might get completely submerged in a bright background if the specimen is more than a several tens of microns thick.

Unlike a standard fluorescence microscope, some microscopes can discriminate the light originating from the focal plane. An ability of a microscope to single out features in the focal volume and filter out the background is called optical sectioning [33]. A number of different optical sectioning microscopy techniques were devised in recent decades, the most widely used of are discussed in the following sections. Light-sheet based fluorescence microscopy (LSFM) and its two implementations that were constructed at EMBL Heidelberg, the *single plane illumination microscope* (SPIM) and the *digital scanned laser light sheet microscope* (DSLM) are more extensively discussed in the Chapter 2.

1.6.2 Confocal microscope

While in an epifluorescence microscope above, the fluorescent specimen is illuminated homogeneously, a laser-scanning confocal microscope (Figure 6) excites only a fraction of the fluorophores using a focused laser beam. An objective lens is usually used to focus the laser beam

⁶ For definitions of numerical contrast, see e.g. *Weber's* and *Michelson's contrast ratios*.

into a small spot inside a specimen. The resulting excitation intensity field near the focus is identical to a wide-field microscope's PSF (Figure 5). An additional mask is inserted in the image plane of the microscope with a small pinhole that transmits light originating from the excitation beam's focal spot but blocks light originating from elsewhere. A photomultiplier tube (PMT) or avalanche photodiode (APD) is used to detect the collected light.

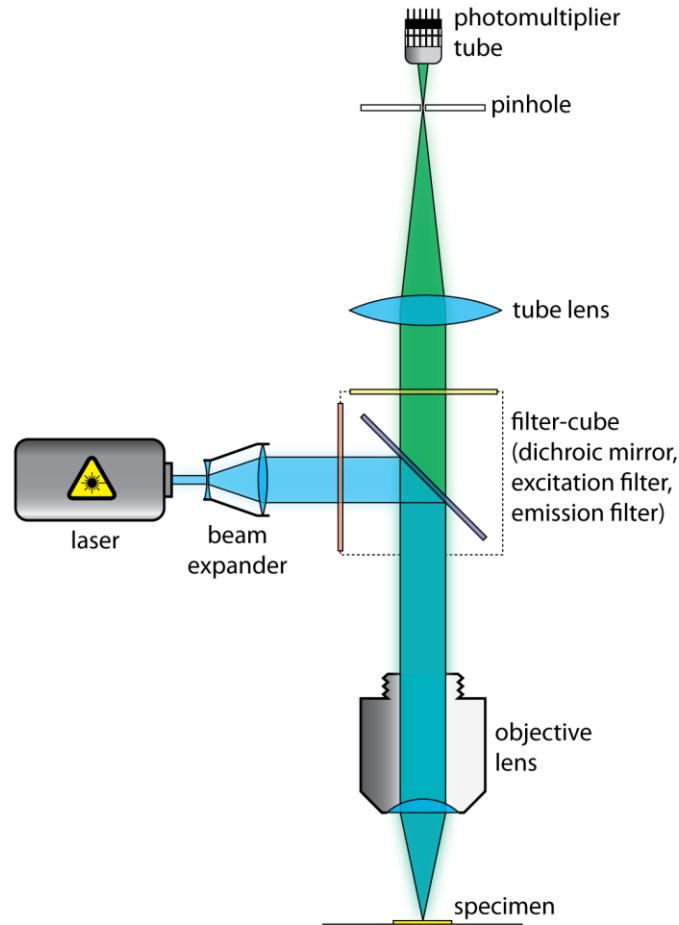


Figure 6: Diagram of a confocal microscope. The essential parts of a confocal fluorescence microscope are shown together with the illumination (blue) and detection (green) pathways. The pinhole in the image-plane of the detection arm discriminates against the light that is not originating from the one single point in the specimen on which the laser beam is focused. The fluorescence light is only measured in a single point. The image is obtained by moving the specimen relative to a stationary beam or by scanning the detection volume relative to a stationary object, which requires additional (not shown) scanning optics behind the objective lens. The confocal fluorescence microscope is a sampling and not an imaging device. To create an image, in most cases the beam is scanned laterally and the object is moved axially.

The probability that fluorescence is detected at a coordinate (x, y, z) now depends on the probability that an excitation light photon is absorbed at this location and on probability that a photon originating from that location is detected by the PMT. The probability that both events take place simultaneously is therefore a product of the probabilities of both individual events. In turn, the resulting system intensity PSF (H_{sys}) of a microscope is calculated as a product of its illumination (H_{ill}) and detection (H_{det}) PSFs:

$$H_{sys}(x, y, z) = H_{ill}(x, y, z) \cdot H_{det}(x, y, z) \quad (1.13)$$

The detection and illumination is normally (but not necessarily, see *confocal theta microscope*, page 25) performed through the same set of optics so the resulting PSF both have a shape of a wide-field microscope's PSF, however each of them corresponding to a different wavelength due to the Stokes shift. Assuming that the axial and lateral PSF profiles are similar to a Gaussian function, the resulting optical resolution (σ_{sys}) can be calculated using the following formula:

$$\frac{1}{\sigma_{sys}^2} = \frac{1}{\sigma_{ill}^2} + \frac{1}{\sigma_{det}^2} \quad (1.14)$$

where σ_{ill} and σ_{det} are optical resolutions of illumination and detection systems, respectively. Both resolutions follow equations (1.2) and the system resolution is therefore

$$\begin{aligned} \sigma_{xy} &= \frac{\lambda_{eff}}{\sqrt{3 - 2 \cos \alpha - \cos 2\alpha}} \\ \sigma_z &= \frac{\lambda_{eff}}{1 - \cos \alpha} \end{aligned} \quad (1.15)$$

where σ_{xy} and σ_z are again lateral and axial resolution, respectively, and λ_{eff} is an effective wavelength calculated by $1/\lambda_{eff} = 1/\lambda_{ill} + 1/\lambda_{det}$. Considering that Stokes shift is usually less than 10% of the detection wavelength, effective wavelength can be sufficiently well approximated by the average wavelength divided by $\sqrt{2}$:

$$\lambda_{eff} = \frac{\bar{\lambda}}{\sqrt{2}} = \frac{\lambda_{ill} + \lambda_{det}}{2\sqrt{2}} \quad (1.16)$$

Resolution of confocal microscope is therefore improved by a factor of $1/\sqrt{2} = 0.71$.

Three examples of a confocal microscope's PSFs (H^{conf}), numerically calculated from relations (1.13) and (1.8), are shown in Figure 7. From the figure it is obvious, that the integral

$$E_{\Sigma}^{conf}(z) = \int_0^{\infty} H^{conf}(r, z) r dr \quad (1.17)$$

is significantly large only near the region around the focal plane ($z = 0$) while it converges to zero with increasing distance from that plane. A confocal microscope therefore is capable of optical sectioning.

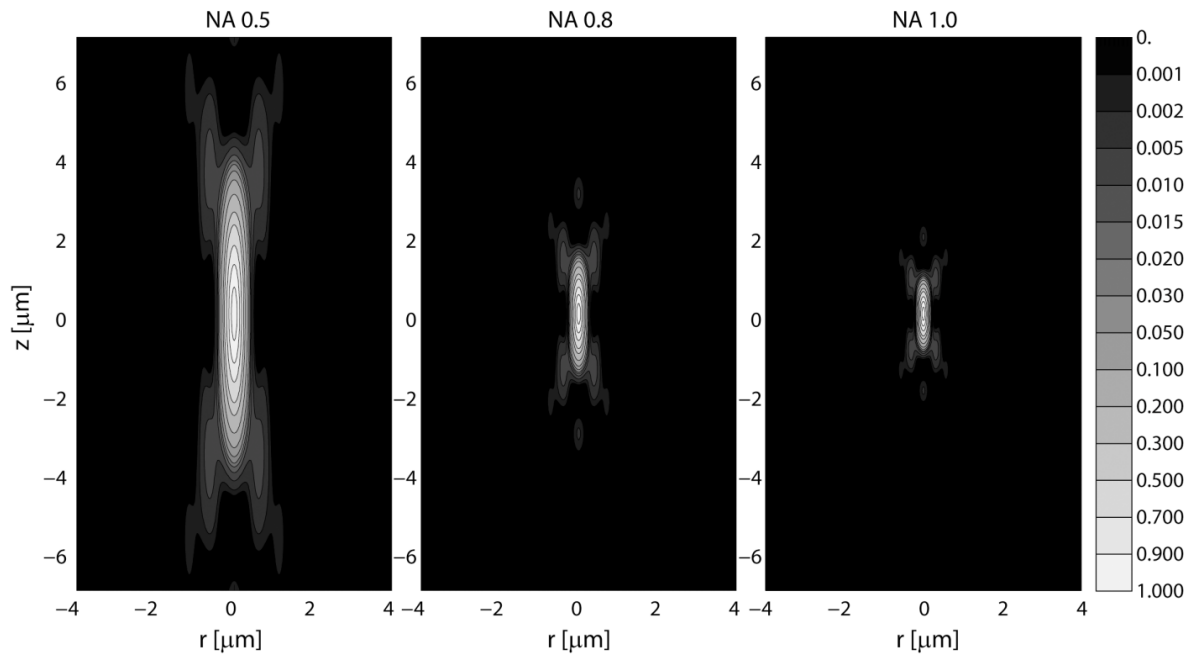


Figure 7: Intensity point spread function of a confocal microscope. The PSF shapes for three different numerical apertures are shown. The calculation (section 1.6.1.1) is based on the Debye diffraction integral [30], water is the immersion medium (refractive index $n = 1.333$) and the wavelength of the light corresponds to the GFP emission peak ($\lambda_0 = 510nm$). The values are maximum-normalized. Again, all three shapes are similar, their extent decreases linearly with the NA along the r -direction and quadratically with the NA along the z -direction with increasing NA. The contour lines were chosen as in [30].

A confocal microscope measures the fluorescence in only one volume element at any time. The three-dimensional image of fluorescence density is created by scanning the specimen through a microscope's detection volume or (more commonly) by scanning the detection volume through the specimen using scanning optics (confocal laser scanning microscope - CLSM).

In a spinning-disk confocal microscope, a mask with a number of pinholes is used to illuminate and collect light from a number of points in the focal plane simultaneously. The mask is rotated at high speed so that the points scan across the field of view and generate an image in the image plane, similarly as a wide field microscope. The image is collected by an image sensor such as a CCD camera. The spinning-disk confocal scheme provides higher frame rates (due to multiple confocal scan volumes) and reduced photo-bleaching compared to CLSM, but provides worse optical sectioning (higher background intensity resulting from out-of-focus fluorescence) and thus worse contrast than CLSM.

1.6.3 Two-photon microscope

This recently developed technique for deep tissue microscopy is based on nonlinear light absorption that occurs at very high light intensities [34,35]. As discussed in section 1.4, a fluorophore's absorption spectrum has a distinctive peak at the wavelength that corresponds to the energy that causes a transition from ground to electronic excited state. However, transition can also take place if two photons with twice the required wavelength are simultaneously absorbed, each of them contributing half of the energy required for the transition, as predicted by Maria Göppert-Mayer in 1931 [36].

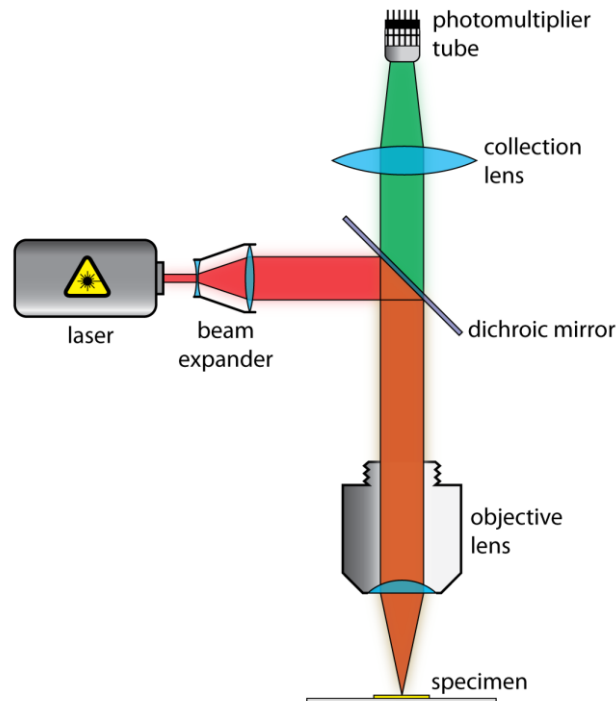


Figure 8: Diagram of a multi-photon epifluorescence microscope. The essential parts of a multi-photon fluorescence microscope are very similar to those in a confocal fluorescence microscope (Figure 7). The pinhole is not required since the optical sectioning is now performed by nonlinear absorption, i.e. in the excitation process. The fluorophores are excited roughly at twice the wavelength light (red light), while the emission light (green light) has the standard spectra of the fluorophore used. Again, fluorescence is measured in one focal volume only. To create an image, the beam is normally scanned laterally and the object is moved axially.

A two-photon absorption requires the simultaneous presence of two photons and its probability is thus proportional to the square of the local light intensity. Two-photon absorption is very unlikely to occur at the light intensities used in common (linear absorption) fluorescence techniques.

Two-photon microscope (Figure 8) is in essence similar to the CLSM microscope: an illumination laser beam is focused into a detection volume through an objective lens and an emitted fluorescence is measured. However, the focal volume is illuminated at twice the wavelength required for the fluorophore's excitation (typically around 1000nm) and very high light intensity. High intensities are reached by use of pulsed lasers that produce very short (down to fractions of a picosecond) light pulses. Focusing is done using high NA lenses (NA 1.0 or above), that produce much smaller focal spots and reach very high light intensities. Due to nonlinear light absorption, most of the emitted fluorescence comes from the focal volume. Even along the optical axis, the fluorescence is constrained to the vicinity of the focal plane, eliminating the need for additional spatial filtering (e.g. confocal pinhole). Two-photon microscope achieves optical sectioning solely by constrained excitation. However, a pinhole can still be added.

Most commonly used fluorophores have absorption peaks in the visible spectrum. The lasers used for two-photon microscopy produce light of twice the wavelengths of the fluorophores' excitation peaks. Typically this refers to the infrared part of the spectrum, from 850nm to 1200nm. However, the spectrum of the emitted fluorescence is still defined exclusively by the fluorophore and is thus normally in the visible part of the spectrum. The excitation and emission spectra are therefore more widely separated than in standard fluorescence techniques. Furthermore, infrared

light has a significantly lower absorption cross-section in biological tissues than visible light and should therefore penetrate deeper (e.g. 500 μm into a mouse brain [37]). The penetration depth is additionally improved by the fact that the spatial discrimination is done solely by means of illumination. The pinhole is not required and all emitted fluorescence can therefore be detected.

Since long wavelength light is used for two-photon microscopy, its resolution is inherently inferior to the other fluorescence microscopy techniques [32]. As mentioned above, the probability of two-photon absorption is proportional to the square of the local light intensity. Two-photon microscope's PSF can therefore be calculated by squaring the $H^{wf}(x, y, z)$ shown in the Figure 5. It generally has the same shape as confocal microscope's PSF (Figure 7) but is larger due to the longer wavelength of the light used.

Other non-linear optics microscopes

Another microscopy technique based on nonlinear light-matter interaction [38] takes advantage of optical high harmonic generation [39]. Second-harmonic generation (SHG) is a process where photon pairs are effectively combined in a material with strong $\chi^{(2)}$ nonlinearity [40] (and thus no inversion symmetry), producing light with double the frequency of the incident light (thus called *frequency doubling*) and approximately the same orientation.

SHG microscopes use pulsed (often femtosecond) infrared lasers for illumination. However, unlike a two-photon microscope, a SHG microscope does not work in an epi- design because the frequency doubled light spreads primarily in the direction of the illumination beam. SHG microscopy therefore requires two opposing sets of optics for illumination and detection. This method also does not involve the use of fluorophores. A strong signal is experienced from highly ordered cellular structures such as lipid membranes, collagen etc.

Another relatively new nonlinear optics based microscopy, coherent anti-Stokes Raman scattering (CARS) microscopy, permits vibrational imaging of specific molecules in unstained samples [41-43]. CARS microscopy has been used in imaging of single lipid bi-layers, polymer films and lipid droplets in adipocyte cells [41].

1.6.4 Other optical sectioning microscopes

Apart from exclusively instrumental methods of achieving optical sectioning discussed above, some sectioning techniques include a certain degree of computational image post-processing.

If the PSF of an imaging system is known, a series of images recorded at different positions of focal plane within the specimen (an image stack) can be computationally deconvolved, effectively inverting the equation (1.5). For an extended review of deconvolution microscopy or computational optical sectioning microscopy (CSOM) see chapter 9 in [44].

Structured illumination microscopy [45] achieves optical sectioning on a standard epifluorescent microscope by spatially modulating the excitation light. The focal plane is illuminated inhomogeneously with a regularly striped pattern. The pattern is then shifted in a step-wise fashion and a set of images (three or more) with different relative positions of the stripes is recorded. A simple digital image fusion is finally used to computationally combine recorded set of images into one image that ideally does not show the modulation of the illumination pattern. Furthermore, as the projected stripes are only sharp in the region close to the focal plane, they are

used to identify the features that lie in that volume. The fusion uses this modulation to discriminate the in-focus features from blurred out-of-focus background.

1.6.5 Lateral vs. axial resolution

As demonstrated in Figure 4, the resolution along the optical axis of a wide-field microscope is considerably worse than the resolution within the focal plane. This manifests itself in a PSF (Figure 5) that is elongated along the optical axis by a factor of at least 2 at high collecting angle objective lenses and significantly more at low aperture lenses.

While the discrepancy proves to be less disturbing when flat specimens (e.g. cells on a slide) are observed and the axial discrimination is not crucial, this poses a serious problem for the three-dimensional microscopy of extended specimens. A confocal microscope theoretically improves both, axial and lateral resolution, by approx. 30%, but the ratio between the two stays the same. Consequently, in the case of non-flat specimens, most microscopes' effective resolving power is limited primarily by its axial resolution.

A number of solutions to mitigate the problem of anisotropic resolution were proposed, most of them employing one or more additional objective lenses. Confocal theta (ϑ) microscopy [46,47] decouples illumination and detection optics and introduces an angle between the both axes (Figure 9). Since PSFs of illumination and detection optics are now elongated along two different directions, the resulting system PSF is more isotropic (Figure 10). The improvement is best when

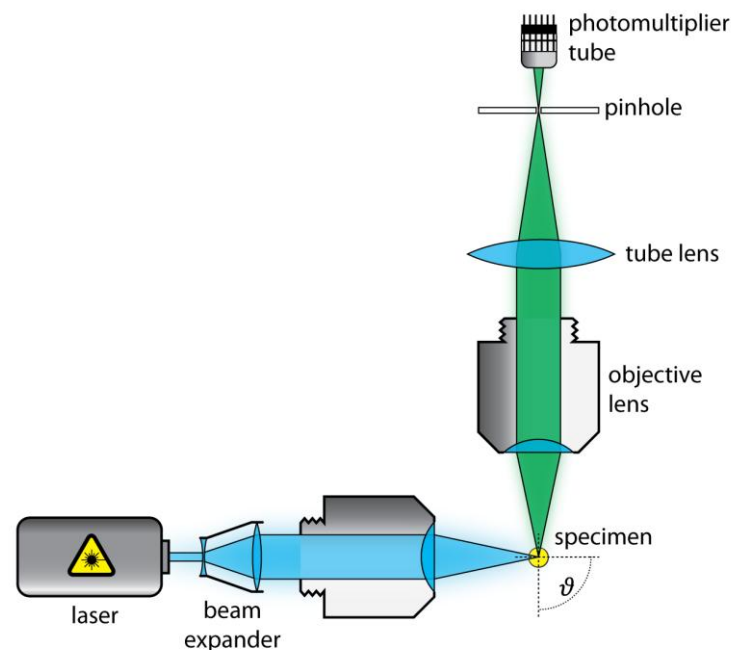


Figure 9: Diagram of a confocal ϑ -fluorescence microscope. The principles of operation are identical to those of a confocal fluorescence microscope (Figure 6). However, the illumination (blue) and detection (green) paths are now decoupled. Both optical axes intersect in the specimen under a certain angle ϑ (here, $\vartheta = 90^\circ$).

the optical axes run normal to each other ($\vartheta = 90^\circ$). The single-lens version of a confocal theta microscope [48,49] employs a small reflective surface near the specimen to reflect the illuminating laser beam into the specimen. The emitted fluorescence is collected directly, i.e. without reflection, so that effective illumination and detection axes are now orthogonal without a need for an additional objective lens.

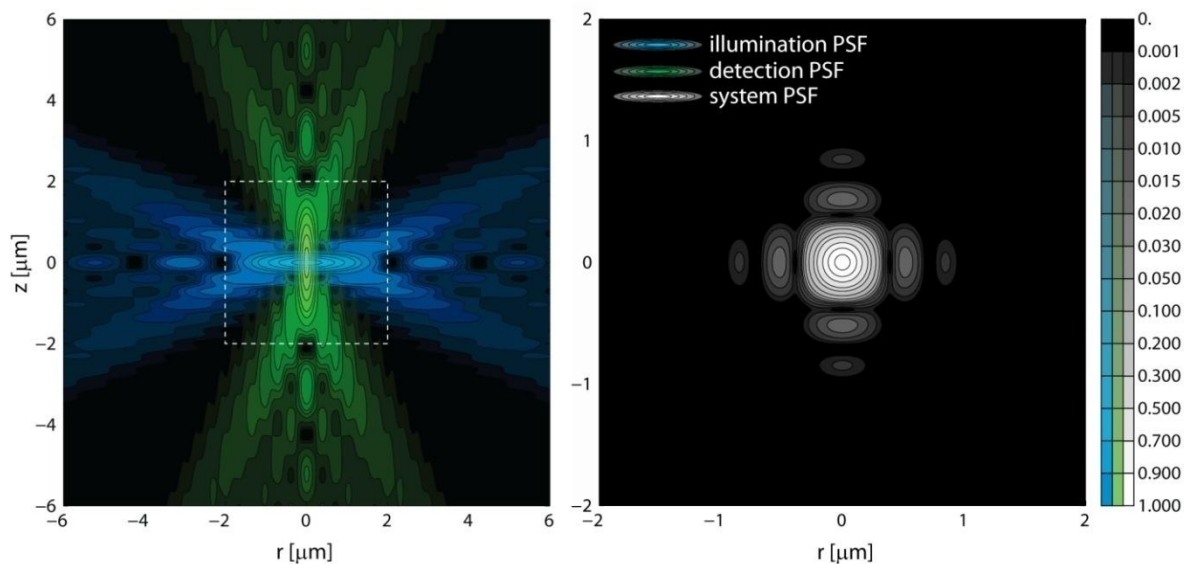


Figure 10: Intensity PSF of a confocal ϑ -fluorescence microscope. Illumination (blue) and detection (green) PSFs have different orientations, producing a system PSF that is more isotropic than that of a confocal fluorescence microscope. The calculation (see section 1.6.1.1) is based on the Debye diffraction integral [30], water as immersion medium (refraction index $n = 1.333$), illumination and detection light wavelengths corresponding to GFP spectrum peaks ($488\text{nm}/510\text{nm}$), $NA = 0.8$ and right angle between the illumination and detection ($\vartheta = 90^\circ$). Values are maximum-normalized. Contour lines were chosen as in [30].

In a standing wave fluorescence microscope (SWFM) [50], two weakly-focused laser beams propagating along opposite directions are used to create a standing wave pattern that illuminates a set of parallel planes that are orthogonal to the optical axis. The illumination pattern is moved axially and a set of images is acquired and computationally fused into an image, with the axial resolution as low as 50nm. The interference pattern is generated by two opposing objective lenses or one lens and a mirror. The method only works well with specimens that are thinner the illumination pattern repeat (approx. 200nm).

Two opposing objective lenses can also be used to effectively double the collecting solid angle. Two distinct realizations were reported: a wide-field Image interference microscope (I^5M) [51] and a point-scanning 4Pi microscope [52]. I^5M microscope combines the standing-wave illumination of the SWFM with bi-directional fluorescence collection. Light collected by the opposing objective lens is coherently combined, forming a wide-field image with an axial resolution below 100nm. The collecting solid angle is still limited to two opposing cones, which cause pronounced side-lobes in axial direction. They are removed computationally by a deconvolution algorithm.

In a 4Pi microscope, the specimen is illuminated coherently through two opposing lenses, illuminating a focal volume with an axial extent of $\sim\lambda/4n$. However, there are strong periodic axial side-lobes with a repeat of $\sim\lambda/2n$. They can be reduced by two-photon excitation [53], theta

detection [54], confocal detection [55], computational deconvolution or by the use of difference between illumination and detection light wavelengths [53]. By combination of 4Pi microscopy with two-photon excitation and confocal detection, up to sevenfold improvement of axial resolution was reported [56].

1.6.6 Super-resolution methods

In 1873 Ernst Abbe faced microscopists with the fact, that the resolution (see section 1.6.1) of an optical microscope is fundamentally limited by the wave nature of the light [57]. Light with wavelengths below 350nm is absorbed by organic tissues and toxic to living cells and thus cannot be used easily to improve the resolution. Additionally, the light collecting half-angle (α in Figure 3) of microscope objective lens is limited to approx. 70° , which corresponds to a numerical aperture of 0.95 in air or 1.45 if immersion oil is used. According to equations (1.2) and (1.3), the best optical resolution attainable by conventional optical microscopy is therefore approx. 250nm laterally and 350nm axially, but even this is practically hardly within reach.

The idea of a near-field microscope, which would use a sub-wavelength sized light sensor a sub-wavelength distance away from a specimen to optically probe specimen's surface with diffraction unlimited resolution, was proposed already in 1928 [58]. The feasibility of the suggestion was first demonstrated with 3cm wavelength microwaves in 1972 [59], before two working examples of a near-field scanning optical microscope (NSOM/SNOM) were simultaneously developed in 1984 [60,61]. Modern SNOM implementations [62] use pointed light-guiding tips to illuminate nanoscopic regions of a specimen's surface and/or to collect the reflected light. The distance of several ten nanometers between the tip probe and the surface is maintained by a feed-back loop. The effective resolutions in the order of $\lambda/25$ are routinely realized and single fluorophore molecules can be observed [63]. However only surfaces can be probed, which makes the technique only marginally useful for life sciences research.

Near-field scanning optical microscope demonstrated, that regions significantly smaller than the wavelength of the light used can indeed be optically probed, if the illumination is constrained to sufficiently small regions. The principle of space-constrained illumination as a means of improving the optical resolution was recently implemented on a wide-field microscope. Gustafsson's structured illumination microscope [64] (not to be confused with Neil's use of the same principle to obtain optical sectioning [45]) illuminates the focal plane with a series of high-frequency periodic patterns with different orientations and phases. A set of wide-field images is recorded and computationally reconstructed into a single image, with a resolution up to two times beyond the Abbe's resolution limit (i.e. 110-120nm [56]).

In the last ten years, new ideas for pushing the far-field fluorescence microscope's resolution to (theoretically) arbitrarily small scales have emerged. This is accomplish by constraining the fluorescence emission to volumes multiple times smaller than the wavelength of the light. Photoactivated localization microscopy (PALM) [65] is based on a novel group of photoactivatable fluorescent proteins that express no fluorescence until they are activated by the illumination with a different (typically shorter) wavelength (section 1.5.1). The activation is a binary, i.e. intrinsically nonlinear process. If the activation light intensity is sufficiently low, only a small fraction of the pool of photoactivatable fluorophores is stochastically activated. The activated fluorescent molecules appear as bright dots in a sufficiently sensitive fluorescence

microscope. The activation light intensity in PALM is normally chosen so low, that the photoactivatable fluorophores are only sparsely activated. Most of the bright dots therefore correspond to single fluorescent molecules. If one can be sure that a bright spot corresponds to a single molecule one can pinpoint its location with a precision of a few tens of nanometers [66]. Apart from the microscope's optical resolution, the localization precision depends on the photochemical stability of the fluorophore and the collection power of the microscope, i.e. how many photons from a fluorophore are detected before it is bleached. Once all of the activated fluorophores are bleached, a new generation is sparsely activated. The procedure is repeated until all the fluorophores have been activated and bleached, i.e. until no new fluorophores can be activated. The procedure provides the positions of individual fluorophores within the specimen with a very high precision. Using the positions of single fluorophores, a synthetic image is then constructed, with an effective resolution reaching down to 20-50nm. Unfortunately, the contrast required for successful localization of single fluorescent molecules necessitates the use of high contrast microscopy methods, such as the total internal reflection fluorescence (TIRF) microscope. This limits PALM's application to thin specimens and no axial discrimination.

A variation of the technique called is stochastic optical reconstruction microscopy (STORM) [67], which uses probes consisting of a photoactivatable reporter fluorophore and an activator that facilitates photo-activation of the reporter molecule in its vicinity. By pairing different (typically cyanine) dyes, many photoactivatable probes with diverse colors were created, paving the way for multicolor high-precision imaging.

A version of the STORM microscope was recently reported that allows high-precision axial localization of single fluorophores [67]. This is realized with a purposely introduced astigmatism in the detection microscope, which asymmetrically deforms microscope's three-dimensional PSF. From the shape of the spot that a single fluorophore creates under such microscope, the fluorophore's location can be pinned down with precision of 50-60nm.

Another recent high-precision method, termed stimulated emission depletion (STED), constrains fluorescence emission volume by means of highly nonlinear emission depletion process [68,69]. The specimen is illuminated with two focused laser beams with different wavelengths. Fluorophores in the beam's focal volume are excited by a short (0.2 ps) pulse with the excitation wavelength, followed by a pulse with a longer STED wavelength. This wavelength's energy corresponds to a fluorophores' transition from the high vibrational/rotational states of the ground electronic level to the bottom of the excited energy band (Figure 1). The STED beam thus stimulates the transition from a highly populated excited electronic state to the ground electronic state with a high vibrational/rotational energy, from where fluorophores rapidly (within picoseconds) relax into the low vibrational/rotational states (see chapter 1.4). The process is thus unidirectional, depleting the pool of excited fluorophores in the path of the STED beam. The STED beam is normally donut-shaped; it has a cavity that overlaps with the focus of the preceding excitation beam. Both, excitation and STED beams follow the Abbe's resolution limit. However, the volume in the center of the STED beam's cavity, where fluorophores are not stimulated to ground state, can have almost arbitrarily small dimensions due to highly nonlinear nature of excitation depletion process. Those fluorophores eventually undergo a normal fluorescent relaxation, producing a light with a shorter wavelength than that of the STED beam and the stimulated fluorescence. Unlike with PALM and STORM above, the position of such sub-

wavelength sized fluorescing volume is now controlled and can be scanned through the specimen. The effective precision depends on the intensity and shape quality of the STED beam. Almost isotropic precisions in the range of 100nm are reported to be routinely achieved while a lateral precision of only 15nm has supposedly already been demonstrated [70].

2 LIGHT-SHEET BASED FLUORESCENCE MICROSCOPE (LSFM)



The next care to be taken, in respect of the Senses, is a supplying of their infirmities with Instruments, and, as it were, the adding of artificial Organs to the natural; this in one of them has been of late years accomplisht with prodigious benefit to all sorts of useful knowledge, by the invention of Optical Glasses. By the means of Telescopes, there is nothing so far distant but may be represented to our view; and by the help of Microscopes, there is nothing so small, as to escape our inquiry; hence there is a new visible World discovered to the understanding. By this means the Heavens are open'd, and a vast number of new Stars, and new Motions, and new Productions appear in them, to which all the ancient Astronomers were utterly Strangers. By this the Earth it self, which lyes so neer us, under our feet, shews quite a new thing to us, and in every little particle of its matter, we now behold almost as great a variety of creatures as we were able before to reckon up on the whole Universe it self.

Robert Hooke

Micrographia, or some Physiological Descriptions of Minute Bodies made by Magnifying Glasses with Observations and Inquiries thereupon (1665), preface

The fundamental idea behind light-sheet based fluorescence microscopy is to combine optical sectioning and wide-field fluorescence microscopy (see section 1.6.1) by illuminating a single, thin section of a fluorescent specimen from the side. The object is thus illuminated orthogonally to the detection optical axis, while the emitted fluorescence light is detected with a standard wide-field fluorescence microscope (Figure 11). This optical arrangement is similar to a confocal ϑ -microscope (see page 25) and this is the reason why LSFM was originally referred to as a wide-field ϑ -fluorescence-microscope [71].

In a properly aligned LSFM, the illuminating light-sheet overlaps with the focal plane of the detection objective lens. The fluorescence emission generated in the light-sheet will therefore originate from a volume close to the focal plane of the detection lens and will form a well focused, high-contrast image.

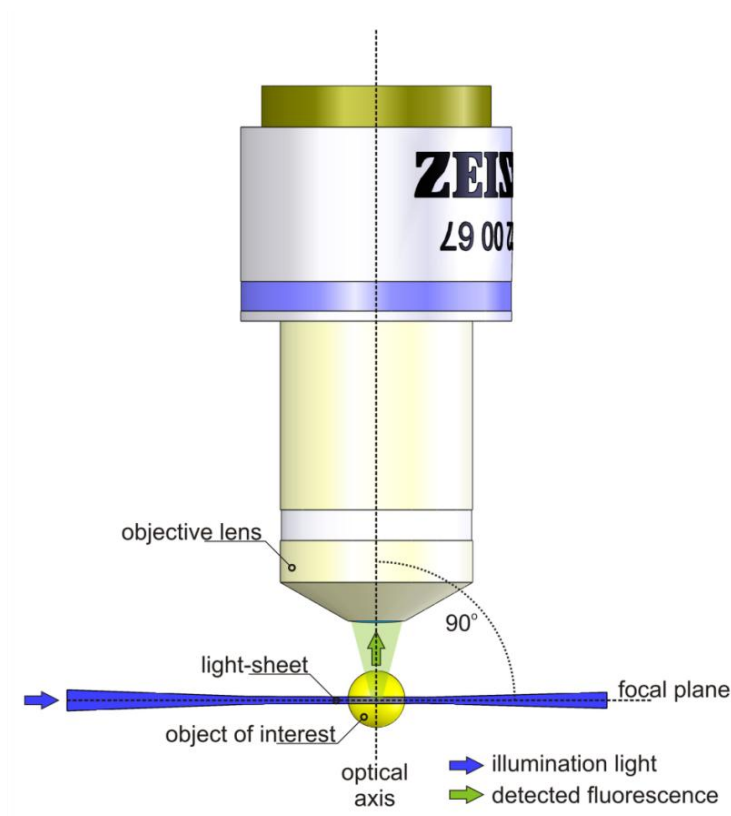


Figure 11: Illumination and detection in LSFM. A thin section is illuminated from the side by a light-sheet (blue), while fluorescent light generated in the fluorescent object is collected normal to the illuminated plane (green). The light-sheet is thinnest in front of the detection objective lens and it overlaps with the objective's focal plane.

Such a microscope bears a number of advantages over other commonly used optical-sectioning microscopes. Since no parts of the specimen outside the light-sheet are illuminated, they are not subject to the undesirable effects of high light intensities, i.e. fluorophore photo-bleaching and more general photo-toxic effects (Figure 12). The reduced per-image photo-damage therefore allows one to increase a recording's duration, dynamic range, sampling rate and frame-rate without increasing the impact on the specimen. Moreover, since one two-dimensional image (i.e. millions of picture elements in parallel) is acquired at a time and since a sensitive CCD camera is used instead of photomultiplier tubes, LSFM produce images with a better signal-to-noise ratio than laser-scanning microscopes.

In this section, the theoretical basis for describing LSFM is presented (section 2.2). Its properties are quantitatively assessed and compared with experimentally determined values. The section also includes technical details of EMBL's LSFMs and general considerations one should be aware of when building an LSFM.

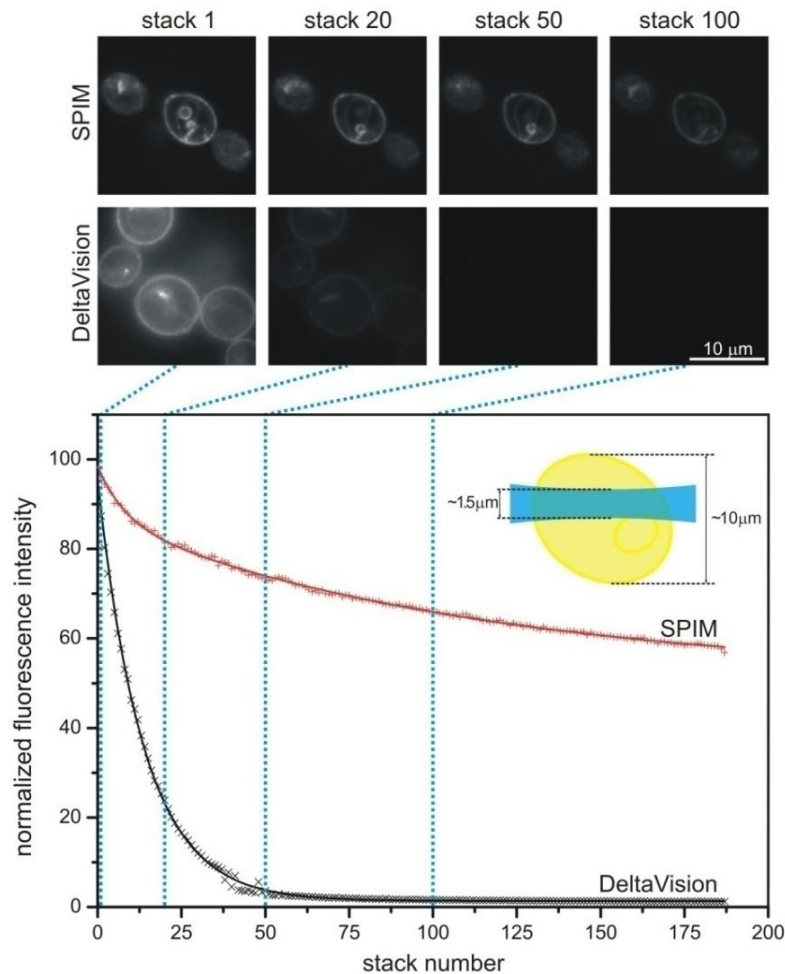


Figure 12: A semi-quantitative comparison of photo-bleaching rates in a SPIM and a regular widefield fluorescence microscope. The yeasts cells stably expressed *Ady2-myeGFP*. A stack of 46 planes was acquired every 8 seconds. A total of 190 stacks were recorded with a SPIM using a Carl Zeiss Achroplan 100x/1.0W lens and an Orca ER CCD camera. The imaging conditions on the two microscopes were adapted to provide comparable signal to noise ratios at comparable excitation intensities. The excitation wavelength was 488nm while the fluorescence emission was recorded above 510nm. The calculation of the fluorescence intensity (crosses in the graph) was performed in the central plane of a yeast cell. The measurements were fitted with a double exponential decay function (solid lines). The fluorescence decay in the widefield microscope was approximately six times faster than in the SPIM. This number is supported by the fact that only one sixth of a yeast cell is illuminated by the light-sheet in SPIM (see the graph inset) while the widefield microscope illuminates the whole cell, i.e. all planes of the stack, for every single image. The difference in bleaching rates is therefore even bigger for larger specimens. It should be stressed that the imaging conditions in such experiments will never be perfect since the sample preparation conditions and samples themselves tend to vary naturally. Published in HFSP Journal [2].

2.1 Use of light-sheets in light microscopy

Light-sheets were and are used in microscopy for more than a century. Heinrich Siedentopf, then a young optical engineer at Carl Zeiss Company in Jena, and the chemist Richard Zsigmondy designed a very early light-sheet based microscope in 1902 to visualize dispersed nano-sized colloidal particles. A thin high-intensity light-sheet was created by illuminating a slit that was demagnified by a microscope objective lens. Light scattered by colloidal particles was then observed by a microscope oriented orthogonal to the illuminated plane [72]. Particles passing through light-sheet appeared as bright spots on dark background. In effect, the device was an

orthogonal-illumination version of a dark-field microscope. The apparatus could visualize particles even when they were considerably smaller than the wavelength of the light used and was therefore named *ultramicroscope*. It enabled a microscopist to determine a sub-wavelength particles' speed, density and estimate their size. It became a fundamental tool for the study of colloids [73]. Richard Zsigmondy was awarded a Nobel Prize for chemistry in 1925 for developing methods that have since become fundamental in colloid chemistry research, including the invention of the ultramicroscope.

Much more recently, the same principle was used for studies of oceanic microbial population distributions [74]. The *Thin Light Sheet Microscope* (TLSM) created a light-sheet by focusing an Argon laser beam by a cylindrical lens. Microscopic organisms traversing the light-sheet were then detected by a wide-field microscope and a CCD camera. As in the ultramicroscope before, the detection of microscopic objects became possible by the high contrast inherent in the orthogonal light-sheet illumination.

A method of photography was patented in 1960, which uses light-sheet illumination to obtain photographs of extended objects in which the whole object is always in focus [75]. It realizes this by illuminating only a section of an object with a light-sheet that is in the focus of the photographic camera. Like in the ultramicroscope, the light-sheet is created by a slit. The object is then slowly moved through the light-sheet while the shutter of the camera is constantly or sequentially opened. Since out-of-focus parts of the object remain dark, they do not contribute to the photograph and sharp images are produced. This illumination method was soon fused with a wide-field microscope, extending the technique to imaging of microscopic objects [76]. Nowadays this popular photographic technique is known as *scanning light macrophotography* [77].

Recently, scanning light macrophotography was extended to produce three-dimensional images. Similarly as in scanning light macrophotography, a technique named *3D light scanning macrography* [78] illuminates only a thin section of an object's surface at a time with a light-sheet. But instead of recording only one image when the whole object is moved through the light-sheet, the object is now translated in equidistant steps and a separate image is recorded at every position. Sections are subsequently reconstructed with a computer to represent the object's three-dimensional surface.

Light-sheet based fluorescence microscope (LSFM) can be regarded as a combination of 3D light scanning macrophotography with fluorescence microscopy (section 1.3). Like in the macrophotography, an object of interest is illuminated by a light-sheet, however in a LSFM only the fluorescent emission from the specimen is collected and used to form an image.

An early reported use of a light-sheet illumination for fluorescence imaging was Orthogonal-plane fluorescence optical sectioning (OPFOS) system that was used to reconstruct a three-dimensional structure of a guinea pig cochlea [79]. Unlike a modern LSFM microscope, OPFOS did not use microscope objective lens and it could not rotate the specimen. The full potential of a LSFM was only unleashed with the Single Plane Illumination Microscope (SPIM) constructed at the European Molecular Biology Laboratory (EMBL) in Heidelberg [71,80,81].

2.2 Basic principles

EMBL's LSFM implementations consist of five basic units:

- a detection unit that collects the fluorescent light, forms an image and records it with a camera,
- an illumination unit, which forms a light-sheet that illuminates a volume that overlaps with the focal plane of the detection lens,
- a specimen translation and rotation unit, which is used to precisely position the specimen relative to the optical system formed by the previous two units,
- a control unit, consisting of a set of electronic controllers and a personal computer running a microscope control software, and
- an offline computer with a software for fusion of multiple-view datasets (section 3).

Additional units can provide extra functionality and can be integrated into a LSFM. The most interesting include laser-ablation [82] (see section 2.4.4), structured illumination [83], fluorescence recovery after photo-bleaching (FRAP) and photo-activation and fluorescence lifetime imaging [84].

2.2.1 Detection unit

The detection unit of a LSFM is basically a wide-field fluorescence microscope (Figure 3) without a dichroic mirror and the illumination optics. It collects light with an objective lens, filters it using spectral filters and forms an image on an image sensor.

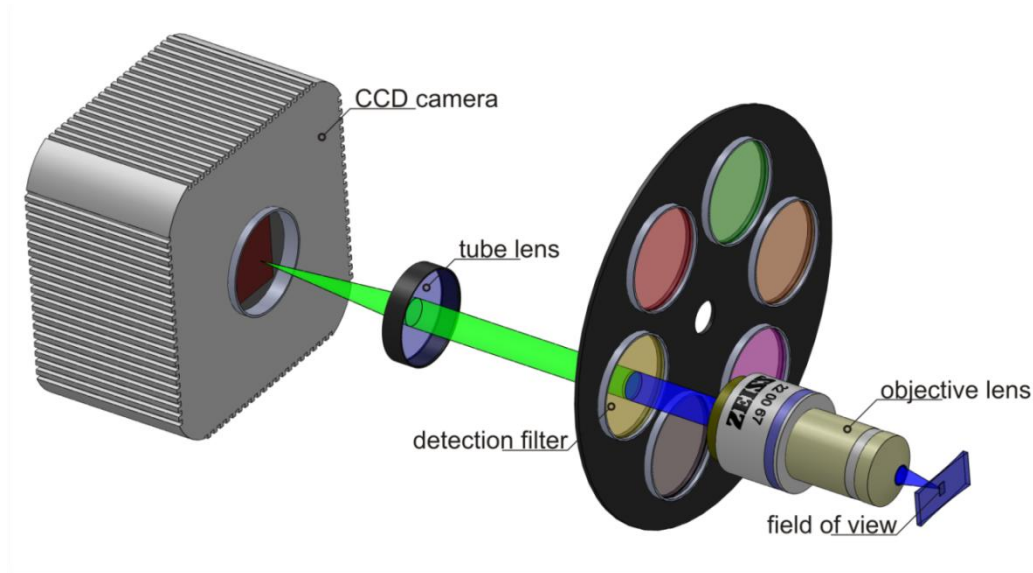


Figure 13: Detection arm of a LSFM. The detection unit of a LSFM is a wide-field fluorescence microscope. It consists of minimally four parts: i) an objective lens, ii) an appropriate detection filter, normally mounted in a filter wheel, iii) a tube lens that in combination with the objective lens form an image on iv) an image sensor (i.e. a camera).

The image sensor consists of an array of light intensity sensitive fields, referred to as picture elements or pixels. Important properties of such a detector are quantum efficiency (i.e. percentage of incident photons that become detected), dynamic range and signal-to-noise ratio (SNR),

detector field size, number of pixels and pixel pitch (distance between centers of neighboring pixels), which is connected to the detector size and number of pixels.

The most commonly used image sensors in biological imaging are based on a CCD (charge-coupled device) chips. They provide quantum efficiencies of up to 70% when front-illuminated and up to 97% in back illuminated electron multiplying CCD (EM-CCD). Recently, CMOS (complementary metal–oxide–semiconductor) detectors are steadily making their way into digital microscopy. They promise frame-rates superior to CCD but are yet to be tested in combination with LSFM.

While cameras with large number of pixels (up to 2048×2048) provide finer sampling and thus improve the resolution of digital images [31] (see also section 1.6.1.1), distributing the same number of photons over a larger number of pixels reduces the SNR and the dynamic range in the images. Detectors with small to moderate number of pixels (up to 1024×1024) are therefore to be preferred for light-critical microscopy. Last but not least, a large number of pixels means more data, which results in slower frame-rates and a more difficult handling of the resulting data (e.g. see *Challenges* on page 71).

The dynamic range is the ratio between the highest and the lowest (non-zero) intensity that can be measured in a single image. The dynamic range is often expressed in bits, i.e. in a base two logarithm of the ratio. Analog-to-digital converters (ADCs) in modern cameras have a dynamic range between 12-16 bits. However, the real dynamic range in an image depends on the SNR and is always lower than the dynamic range of the camera's ADC. In LSFM, noise is dominated by the Poisson statistics resulting from the light quantization.

All current EMBL's LSFM implementations are based on two CCD cameras:

	Hamamatsu ORCA-AG	pco.imaging pco.2000
Sensor size (diagonal)	8.67 mm × 6.6 mm (10.9 mm)	15.6 mm × 15.3 mm (21.9 mm)
Pixel resolution (pitch)	1344 × 1024 (6.45 μm)	2112 × 2072 (7.4 μm)
Quantum efficiency	400 nm	52 %
	500 nm	70 %
	600 nm	64 %
	700 nm	47 %
	max sensitivity	72 % @ 555 nm
ADC dynamic range	12 bit	14 bit
Full sensor frame rate	8.8 frames/s (one ADC)	14.7 frames/s (two ADCs)

Objective lens

The objective lens collects the light emanating from its field of view. The traditional characteristics of microscope objective lenses are the magnification (M), the numerical aperture (NA), the immersion medium's refractive index and the working distance. The magnification is the ratio of the size of an object and the size of its image that is formed in combination with a prescribed tube lens (e.g. virtually all Carl Zeiss objective lens are designed for a tube lens with a focal length of 164.5mm). The choice of the magnification is a matter of a tradeoff between the resolution (defined by objective's NA, see page 14) and the size of the field of view. High magnification objective lenses have generally a higher NA (better resolution) but image a smaller field of view. The latter, however, also depends on the size of the image sensor.

Currently, LSFM are mostly used with biological specimens, which usually require water based media. This is why water dipping lenses are popular with LSFM. A dipping lens removes the need for additional air/water/glass interfaces between the objective lens and the object, but requires the use of medium filled chambers that enclose the objective front lens and the specimen (Figure 14 and Figure 15). Water dipping lenses used in LSFM have smaller NAs (up to 1.1) than oil immersion lenses (up to 1.45), but also relatively long working distances, which makes specimen handling during imaging considerably easier. Most low magnification, low NA objective lenses (5x and less) are designed for use without an immersion medium. However, water based gels, normally used for LSFM specimen preparation (section 2.3), have to be immersed in an aqueous medium. Air lenses are therefore frequently used in combination with glass-walled imaging chambers that contain the medium and allow imaging through its transparent walls. It has to be noted that air objective lenses were not designed to be used like that. Use of objective lens in modes that the lens was not designed for, often results in an effective magnification and working distance that are considerably different from those specified on the objective lenses. Image aberrations (especially chromatic and spherical aberrations) might be significantly increased as well.

EMBL's LSFM microscopes are used with the following set of detection objective lenses (note that field of view depends on the camera's sensor size):

Lens (Carl Zeiss)	Immersion medium	Resolution [μm] ⁷		Field of view dimensions	
		Lateral	Axial	Orca-AG	pco.2000
Fluar, 2.5x/0.12	Air	2.448	93.832	3.5mm × 2.6mm	6.2mm × 6.1mm
Fluar, 5x/0.25	Air	1.174	21.470	1.7mm × 1.3mm	3.1mm × 3.1mm
Achroplan, 10x/0.3W	Water	0.977	14.850	867 μm × 660 μm	1.6mm × 1.5mm
Achroplan, 20x/0.5W	Water	0.584	5.217	434 μm × 330 μm	780 μm × 765 μm
Plan-Apochromat, 20x/1.0W	Water	0.284	1.123	434 μm × 330 μm	780 μm × 765 μm
Achroplan, 40x/0.8W	Water	0.361	1.903	217 μm × 165 μm	390 μm × 383 μm
Achroplan, 63x/0.9W	Water	0.319	1.451	138 μm × 105 μm	248 μm × 243 μm
Achroplan, 100x/1.0W	Water	0.284	1.123	87 μm × 66 μm	156 μm × 153 μm

The magnifications stated on the objective lenses require a combination with an appropriate (164.5 mm focal distance) tube lenses. Tube lenses with different focal lengths can be used to tune the magnification to the size of the objects without changing the NA. Alternatively, one can alter the total magnification by inserting a telescope element (available commercially e.g. by Carl Zeiss as n *OptoVar*) between the objective lens and the tube lens.

⁷ Calculated according to SGH method [28] (see section 1.6.1)

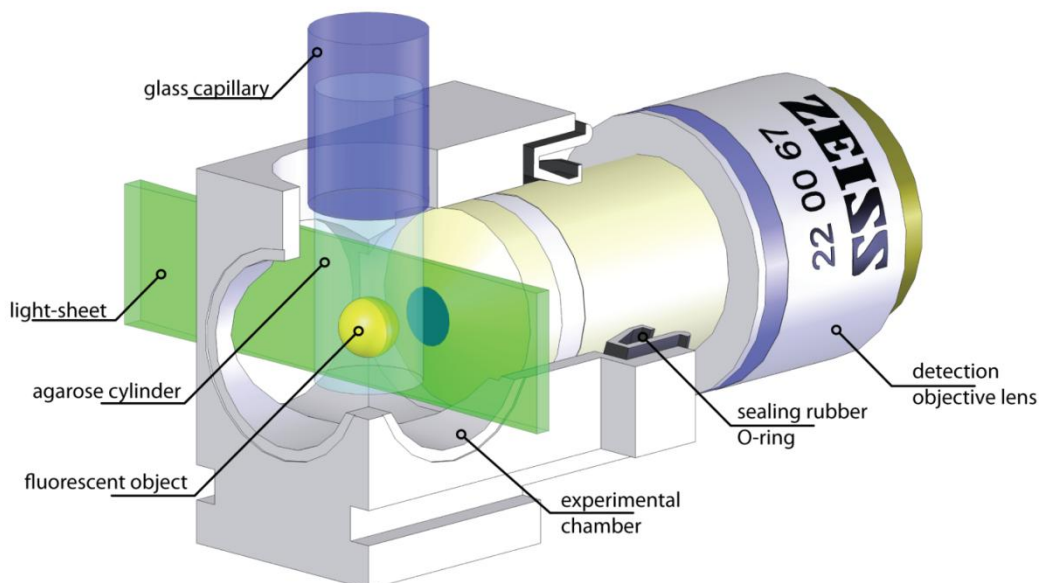


Figure 14: Experimental chamber of a LSFM. LSFM is used mostly with water-dipping lenses. The specimen is immersed into a water-based medium contained in a chamber. The chamber has three glass windows, one of them being used for the light-sheet illumination, and the detection objective lens is inserted through a sealing rubber O-ring (shaft seal) in the remaining side. Size of the light sheet is exaggerated.

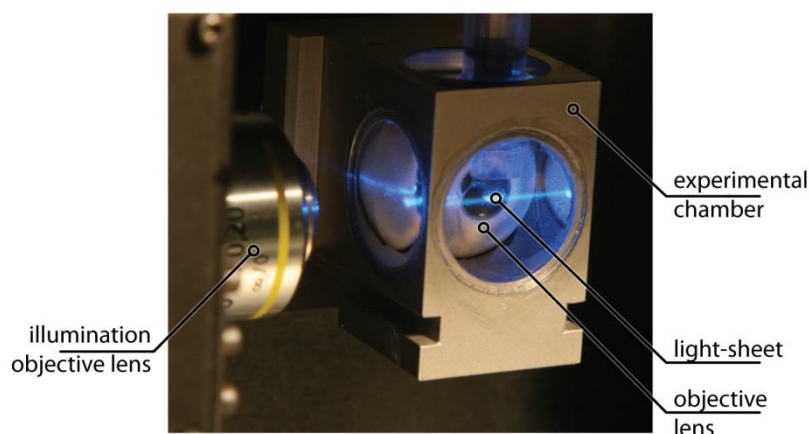


Figure 15: Photograph of a LSFM experimental chamber. The blue light-sheet, projected from the left, can be seen since the light is scattered in the water and by the agarose. The detection objective lens is inserted from the back. The bottom edge of the chamber is 20mm long.

2.2.1.1 Emission filter

The emission filter blocks the light that is scattered in the specimen and transmits the fluorescence emission. The latter has a longer wavelength due to the Stokes shift (see section 1.5) therefore long-pass filters are commonly used. The filters are mounted in a filter wheel that rapidly (within less than 40ms) changes filters when multiple fluorescence channels (different fluorescent dyes) are imaged.

The long-pass filters transmit fluorescence light more efficiently than band-pass filters. The cut-off frequency should be slightly above the illumination wavelength (e.g. the company Semrock, Rochester, New York, USA, offers filters with transition widths as low as 3-6 nm). The band-pass

filters are useful when selectivity is more critical than light efficiency. For example, nonspecific autofluorescence has a broader spectrum than commonly used fluorophores. Therefore, the signal intensity relative to the autofluorescent background is significantly improved by the use of an appropriate band-pass filter. This is especially important when imaging short wave-length fluorophores (DAPI, GFP), where the autofluorescent absorption is relatively high. The band-pass filters are also helpful when specimens contain multiple different fluorophores with similar spectra that cause inter-channel “cross-bleeding”.

2.2.2 Illumination unit

The illumination distinguishes an LSFM from a wide-field fluorescence microscope. Good optical slicing and excellent contrast are only achieved when a very thin section of a specimen is illuminated. A typical light-sheet in LSFM is only 1.5 to 4 μm wide, i.e. only a couple of light wavelengths. The fundamental limit to the thickness of the light-sheet and the resulting axial resolution of an LSFM stems from the wave nature of the light propagation. The following discussion is, therefore, based on the theory of diffraction.

Lasers are common light sources in modern fluorescence microscopy. For paraxial waves, i.e. waves with nearly parallel wave front normals, a general wave equation can be approximated with the paraxial Helmholtz equation [40,85]:

$$\nabla_{\perp}^2 U + i 2k \frac{\partial U}{\partial z} = 0 \quad (2.1)$$

Without loss of generality we assume that the light spreads along the \hat{z} axis. Furthermore, $\nabla_{\perp}^2 = \partial^2/\partial x^2 + \partial^2/\partial y^2$ is the transverse Laplace operator, $U(\vec{r})$ is the wave-function, $k = 2\pi/\lambda$ is referred to as wave number and $\lambda = \lambda_0/n$ is the wavelength of the light in a medium with a refraction index n .

A simple Eigen-function of the partial differential equation (2.1) is a Gaussian beam:

$$U(r, z) = A_0 \cdot \frac{W_0}{W(z)} \cdot e^{-\frac{r^2}{W^2(z)}} \cdot e^{-i\varphi(r, z)} \quad (2.2)$$

Where A_0 defines the amplitude of the wave, W_0 is the radius of the beam at its thinnest location (beam waist), $r = \sqrt{x^2 + y^2}$ is the distance from the beam's axis, $\varphi(r, z)$ is the combined phase part of the wave-function and $W(z)$ is the radius of the beam at a distance z from the waist and has a hyperbolic shape:

$$W(z) = W_0 \sqrt{1 + \left(\frac{z}{z_0}\right)^2} \quad (2.3)$$

The parameter z_0 in equation (2.3) is called Rayleigh range. The Rayleigh range and the beam waist radius W_0 are connected by the following relation:

$$z_0 = \frac{\pi W_0^2}{\lambda} \quad (2.4)$$

A beam with a small waist radius therefore has a short Rayleigh range, i.e. it is thin in a shorter region around its waist, and it expands faster outside. Far from the waist, the beam divergence (2.3) becomes linear:

$$W(z) \approx \alpha \cdot z \quad \text{when } z \gg z_0 \quad (2.5)$$

where α is the half-angle of the beam divergence

$$\alpha = \frac{W_0}{z_0} = \frac{\lambda}{2W_0} \quad (2.6)$$

The intensity of the emitted fluorescence is determined by the intensity of excitation light. The intensity of a Gaussian beam (2.2) (1.2) is

$$I(r, z) = U(r, z) \cdot U^*(r, z) = |A_0|^2 \cdot \left(\frac{W_0}{W(z)}\right)^2 \cdot e^{-\frac{2r^2}{W^2(z)}} \quad (2.7)$$

Another solution of the paraxial Helmholtz equation (2.1) is an elliptical Gaussian beam:

$$U(x, y, z) = A_0 \cdot \sqrt{\frac{W_{x,0}}{W_x(z)}} \sqrt{\frac{W_{y,0}}{W_y(z)}} \cdot e^{-\frac{x^2}{W_x^2(z)}} \cdot e^{-\frac{y^2}{W_y^2(z)}} \cdot e^{-i \cdot \varphi(x,y,z)} \quad (2.8)$$

This beam has a Gaussian profile along the \hat{x} and \hat{y} axes, but the two waist radii are uncoupled. Beam thicknesses and Rayleigh ranges along both axes can now be different:

$$\begin{aligned} z_{x,0} &= \frac{\pi W_{x,0}^2}{\lambda} \\ z_{y,0} &= \frac{\pi W_{y,0}^2}{\lambda} \end{aligned} \quad (2.9)$$

The light intensity in the beam is, therefore,

$$I(x, y, z) = U(x, y, z) \cdot U^*(x, y, z) = |A_0|^2 \cdot \frac{W_{x,0}}{W_x(z)} \cdot \frac{W_{y,0}}{W_y(z)} \cdot e^{-\frac{2x^2}{W_x^2(z)}} \cdot e^{-\frac{2y^2}{W_y^2(z)}} \quad (2.10)$$

where, again

$$W_x(z) = W_{x,0} \sqrt{1 + \left(\frac{z}{z_{x,0}}\right)^2} \quad \text{and} \quad W_y(z) = W_{y,0} \sqrt{1 + \left(\frac{z}{z_{y,0}}\right)^2} \quad (2.11)$$

Such elliptic beams are used in SPIM for planar specimen illumination. The ratio between the beam extensions along both dimensions $W_{x,0}/W_{y,0}$ is around 200 at low magnifications (2.5x-10x) and as low as 40 at high magnifications (100x).

2.2.3 Single plane illumination microscope (SPIM)

In EMBL's first implementation of LSFM [71] a light-sheet illumination was created using a cylindrical lens that focused a laser beam along one dimension (\hat{y} in Figure 16a) while keeping it wide and thus collimated along the other (\hat{x} in Figure 16a). The light-sheet, i.e. an elliptical Gaussian beam created by such a system, was tuned to the size of the field of view of the detection system by a mask (a slit).

2.2.3.1 Light-sheet thickness

A thick light sheet provides suboptimal slicing and contrast in the center of (and potentially across the whole) field of view (FOV). On the other hand, from (2.9) it follows that a thin light-sheet corresponds to a short Rayleigh length, which might result in a thick light-sheet at the edges of the FOV and suboptimal sectioning. The light-sheet thickness must therefore be adjusted to the extent of the FOV along the illumination axis. As a rule of thumb, we try to make the edges of the FOV correspond to the edges of the waist region of the light-sheet (Figure 16). This means that the edges of the FOV are approximately one Rayleigh distance away from its center; field of view extent along the illumination axis equals two Rayleigh lengths: $W_{\text{fov}} = 2Z_{y,0}$. According to formula (2.11), the light-sheet at the edges will be approximately 41% thicker than in the center. Such variance across the FOV is very much tolerable for most of the imaging applications.

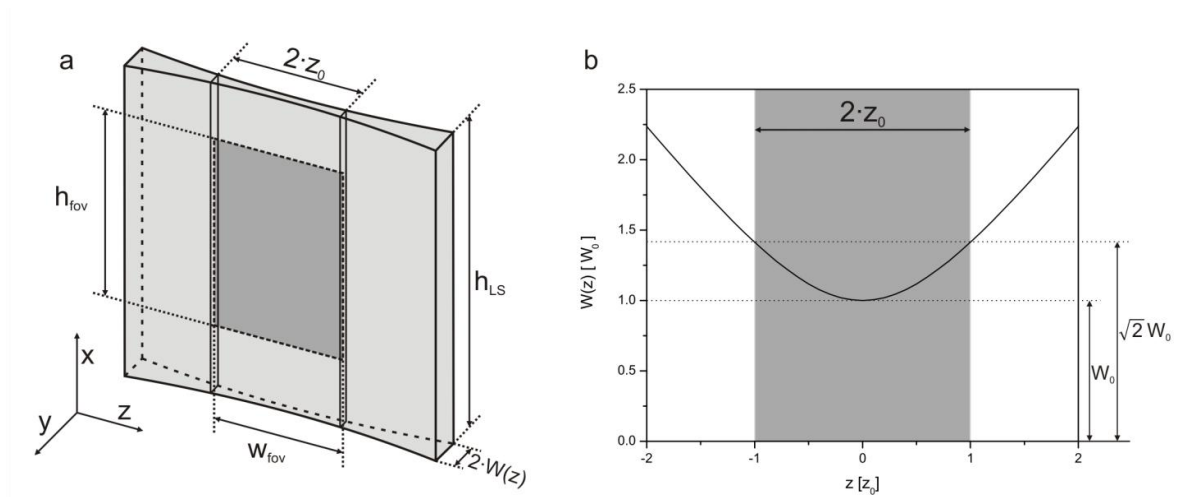


Figure 16: Light-sheet dimensions. The shape of the light-sheet along the direction of the illumination (axis z in a) is hyperbolic. It is thin in the middle of the field of view (FOV) and widens towards the edges (a). The thickness of the thinnest part (i.e. the beam waist) of the light-sheet is chosen such that the boundaries of the FOV lie approximately one Rayleigh-range distance from the waist. This means that the light sheet at the edges of the FOV is approx. $\sqrt{2} \cong 1.4$ times thicker than in the middle (b).

Quantitatively, the light-sheet thickness (LST) is defined as the full width of the Gaussian profile along \hat{y} in (2.10) :

$$LST = 2\sigma(z) = W_y(z) \quad (2.12)$$

From (2.10) it follows that the volume inside the light-sheet thickness is traversed by 68% of the total illumination power.

If the FOV is not square, it is best to orient the image sensor such that the short edge of the FOV (W_{fov} in Figure 16) is parallel to the direction of the illumination. From equation (2.9) and $W_{fov} = 2Z_{y,0}$ we can calculate the following relation between the light-sheet waist thickness and the size of the FOV along the illumination axis:

$$LST_0 = W_{y,0} = \sqrt{\frac{\lambda}{2\pi} \cdot W_{fov}} = \sqrt{\frac{\lambda}{2\pi} \cdot \frac{W_{CCD}}{M}} \quad (2.13)$$

where W_{CCD} stands for the width of the image sensor along the illumination axis and M is magnification of the detection optics.

The optimal angular aperture of the light-sheet follows from equations (2.13) and (2.6) :

$$\alpha = \sqrt{\frac{\lambda \pi}{2 W_{fov}}} = \sqrt{\frac{\lambda \pi}{2} \cdot \frac{M}{W_{CCD}}} \quad (2.14)$$

The angular aperture of the illumination beam remains relatively small (e.g. $0.11\text{rad} = 6.2^\circ$ even at 100x magnification and Orca camera). The assumptions for the paraxial approximation, which this discussion of the illumination optics is based on, are therefore valid.

2.2.3.2 Light-sheet height

The light-sheet has a Gaussian profile along both lateral directions (\hat{x} and \hat{y} in the calculations above). However, it is thinner across the light-sheet (\hat{y}) while it has to be wide enough along the light-sheet (\hat{x}) to illuminate the whole field of view as evenly as possible. Only the area around the centre of the Gaussian profile is used, where the intensity does not vary as much (Figure 17, left). This, in turn, means that a part of the total illumination power is wasted on the areas outside the field of view (or blocked by a mask in the illumination optics, placed in a plane conjugate to the plane of the waist).

Let us define two variables: an illumination uniformity u , which measures the illumination intensity at the edge of the field of view as a fraction of the maximum intensity in the centre of the field of view, and a power utilization t that refers to what fraction of the total illumination power traverses the field of view of the detection optics and might result in detected fluorescence.

From (2.10) it follows that $u = e^{-\frac{H_{fov}^2}{W_{\hat{x}}^2}}$, where H_{fov} is height of the field of view. If we integrate eq. (2.10) along \hat{x} from $-\frac{H_{fov}}{2}$ to $\frac{H_{fov}}{2}$ we obtain the following relation between uniformity and power utilization

$$t = \text{erf}\left(\sqrt{\ln \frac{1}{u}}\right) \quad (2.15)$$

where $\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$ is the Gauss error function. The relation demonstrates the tradeoff between uniformity and power utilization (Figure 17, right). The uniformity is near unity only when a small fraction of the total power is utilized and most of the light is wasted. In a real

application, the uniformity usually lies somewhere around 0.8 (20% reduction of the intensity at the image edges), which means that around half of the total illumination power is utilized (dashed line in Figure 17, right). Considering the power of modern lasers, this fraction is more than sufficient.

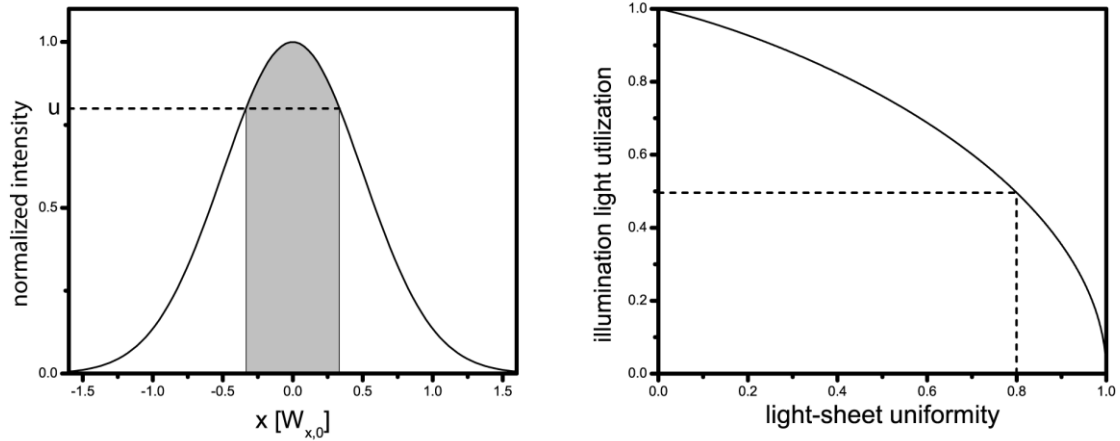


Figure 17: Light-sheet height. A SPIM's light-sheet has a Gaussian profile along its two lateral principal directions. The width of the profile along the light-sheet has to be wide enough to illuminate the whole field of view as evenly as required. Only the region around the maximal intensity is therefore used (gray area in the left graph), while the rest is discarded. The ratio between the illumination intensity in the centre and at the edges of the field of view, i.e. illumination uniformity (u in the left graph) will be higher, if more illumination power is wasted (low utilization). The tradeoff between the illumination power utilization and the light-sheet uniformity is shown in the right graph. The dashed line marks the properties of a commonly used light-sheet profile.

2.2.3.3 Resolution

SPIM's detection is based on a regular wide-field microscope (section 1.6.1). The planar illumination has no effect on the lateral resolution of the SPIM, which is thus described by the formula (1.2) on page 15. Axial resolution, on the other hand, is simultaneously governed by axial discrimination of the detection optics and the single plane illumination.

SGH theory predicts the following lateral (σ_{xy}) and axial (σ_z) resolutions of a cylindrical lens with the angular aperture α [28]:

$$\begin{aligned}\sigma_{xy} &= \frac{\lambda}{2} \sqrt{\frac{\alpha}{2\alpha - \sin 2\alpha}} \\ \sigma_z &= \frac{\lambda}{2} \frac{\alpha}{\sqrt{2\alpha^2 - 4\sin^2 \alpha + \alpha \sin 2\alpha}}\end{aligned}\tag{2.16}$$

Since the axial direction of SPIM corresponds to the lateral direction of its illumination optics, the axial resolution due to single plane illumination corresponds to σ_{xy} . The angular aperture α is usually small (see section 2.2.3.1) so the first formula in (2.16) can be approximated linearly by $\sigma_{xy} = \frac{\sqrt{3}\lambda}{4\alpha}$. Adding relation (2.14), the following illumination part of the axial resolution in SPIM is obtained:

$$\sigma_{z,ill} = \sqrt{\frac{3 \lambda W_{fov}}{8 \pi}} \quad (2.17)$$

In the context of SGH theory, the detection part of the SPIM's axial resolution is expressed by the formula (1.3) on page 15:

$$\sigma_{z,det} = \frac{\lambda}{1 - \cos \alpha} \quad (2.18)$$

Again, the combined axial resolution is calculated in the inverse-Pythagorean manner, equation (1.14) on page 21:

$$\sigma_{sys} = \frac{1}{\sqrt{1/\sigma_{ill}^2 + 1/\sigma_{det}^2}} \quad (2.19)$$

The total axial resolution of a SPIM microscope (σ_{sys}) is therefore always better than the illumination and detection axial resolutions. If σ_{ill} and σ_{det} are considerably different, σ_{sys} is approximately equal to the smaller of them; if they are similar (as in the case of a confocal microscope), σ_{sys} is smaller by approximately a factor of $\sqrt{2}$. The axial resolution improvement due to a planar illumination for the commonly used microscope objective lens is demonstrated in the Figure 18 and in the following table:

Lens (Carl Zeiss)	Detection resolution			Light-sheet thickness [μm]	Total axial resolution [μm]	Axial resolution gain	PSF elongatio n
	Lateral [μm]	Axial [μm]	Ratio				
Fluar, 2.5x/0.12	2.448	93.832	38.3	14.31	12.29	7.63	3.12
Fluar, 5x/0.25	1.174	21.470	18.3	10.12	8.12	2.64	2.25
Achroplan, 10x/0.3W	0.977	14.850	15.2	7.16	5.72	2.60	2.66
Achroplan, 20x/0.5W	0.584	5.217	8.9	5.06	3.36	1.55	2.66
Plan-Apochromat, 20x/1.0W	0.284	1.123	4.0	5.06	1.09	1.03	3.63
Achroplan, 40x/0.8W	0.361	1.903	5.3	3.58	1.62	1.17	3.25
Achroplan, 63x/0.9W	0.319	1.451	4.5	2.85	1.25	1.16	3.64
Achroplan, 100x/1.0W	0.284	1.123	4.0	2.26	0.97	1.15	4.05

The values were calculated based on the following assumptions: 488nm illumination light, 510nm detection light, Orca ER image sensor (detector width 6.6mm).

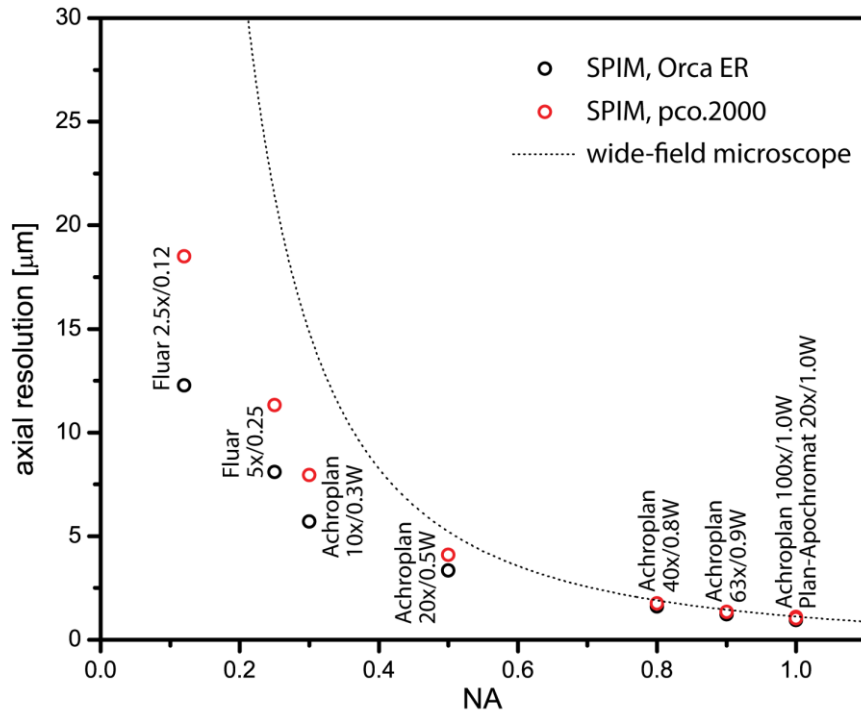


Figure 18: The axial resolution of a LSFM. The axial resolution in LSFM is influenced primarily by the axial resolution of the detection lens at high NAs (>0.5), while at low NAs, the planar illumination prevents a rapid decrease of the axial resolution characteristic to wide-field microscopes (dashed line). Axial resolution at low NAs depends on thickness of the light-sheet and consequently on the size of the field of view. The circles in the graph represent axial resolution of LSFM for different combinations of common objective lens and cameras.

The axial resolution is most dramatically improved at low NA lens (0.5 and below), while the enhancement remains significant even at highest magnifications (15% at 100x, NA 1.0). Consequently, the PSF elongation (ratio between axial and lateral resolution) is now only weakly influenced by the NA of the detection lens. In LSFM it stays low even at low NAs, while it diverges in a wide-field microscope (see Figure 4). For example, the PSF elongation in LSFM is 3.12 at NA 0.12 in the table above, but 38.3 in a wide-field microscope. It should be emphasized that even when the resolution improvement due to planar illumination is low (i.e. at high magnifications), SPIM provides optical sectioning, which a wide-field microscope lacks.

The shape of the PSF can be calculated by multiplying the PSFs of illumination and detection optics. The former is the solution of the Debye's diffraction integral (1.6) and the later is given by the equation (2.10). The shapes of both contributions and their product for three objective lenses commonly used with EMBL's LSFM implementations are illustrated in Figure 19.

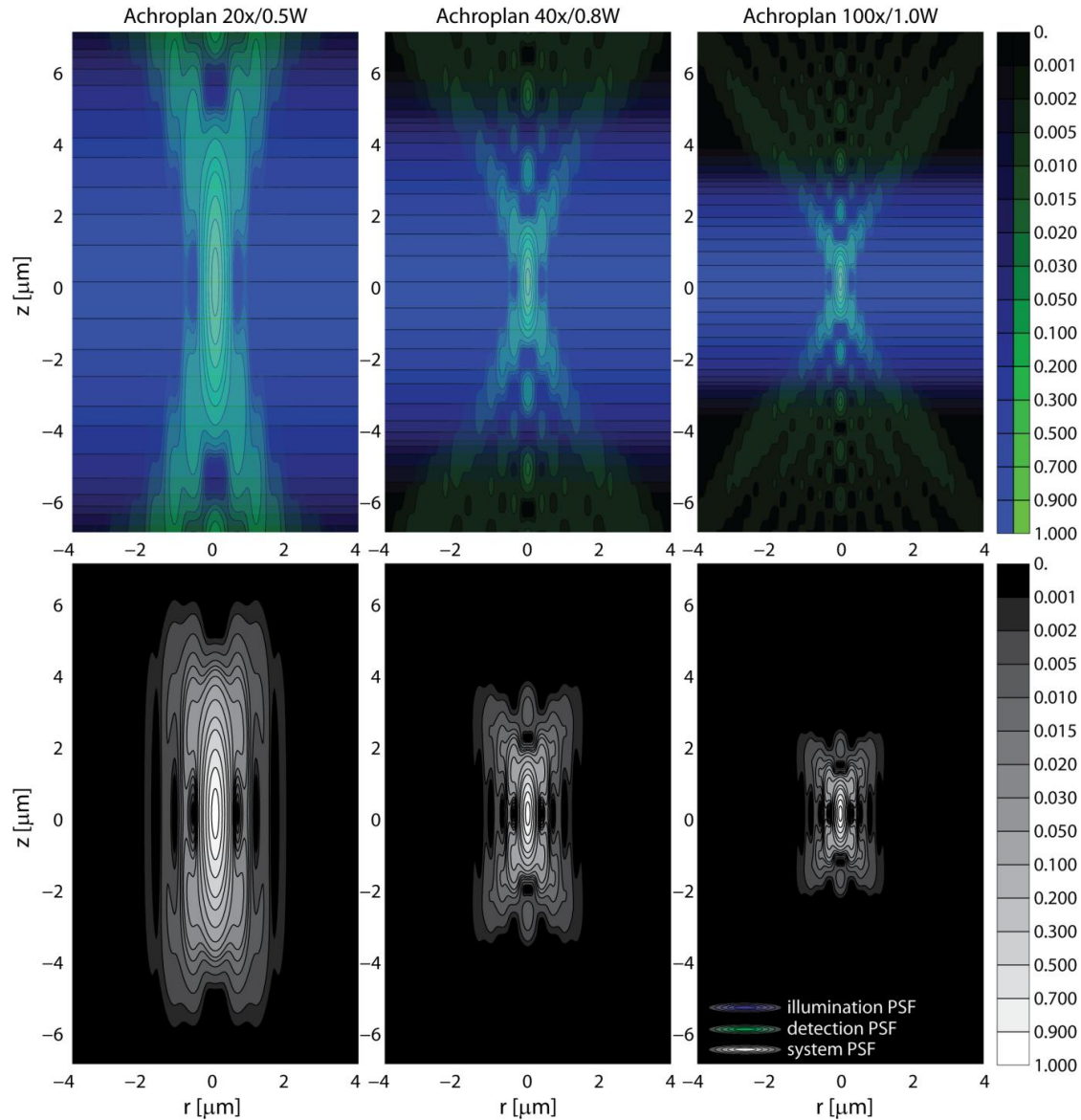


Figure 19: Intensity PSF of a LSFM for three common detection objective lenses. The point spread function of a LSFM is defined by a combination of its illumination (blue) and detection (green) PSFs. The illumination PSF has a Gaussian profile along z while the detection PSF is the same as that of a wide-field microscope (Figure 5). The system PSF of LSFM can be seen in the bottom row. Note how the first minimum along z disappears at NAs of 0.5 and below, where the axial resolution of the microscope is improved by the planar illumination.

2.2.3.4 Contrast improvement

The light-sheet illuminates only a thin section of an object. The light-sheet intensity profile (2.10) across the focal plane is Gaussian and, therefore, theoretically, nonzero everywhere. However, the bulk of the total power is concentrated in a thin volume around the focal plane. The fluorescence is generated predominantly in this thin section of the specimen. Compared to a fluorescence wide-field microscope, where the whole specimen is illuminated, SPIM images therefore have an improved contrast, i.e. an intensity of the in-focus signal over the out-of-focus background.

Light-sheet illumination improves the contrast whenever the specimen is larger than the light-sheet. However, due to the finite thickness of the light-sheet, not all emitted fluorescence comes

from the focal volume of the detection optics. The fraction of the total power that illuminates the in-focus part of the specimen depends on the ratio between the detection unit's depth of field (DoF) [30] and the thickness of the light-sheet:

$$p = \operatorname{erf}\left(\frac{1}{\sqrt{2}} \frac{\text{DoF}}{\text{LST}}\right) \quad (2.20)$$

where $\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$ is the Gauss error function.

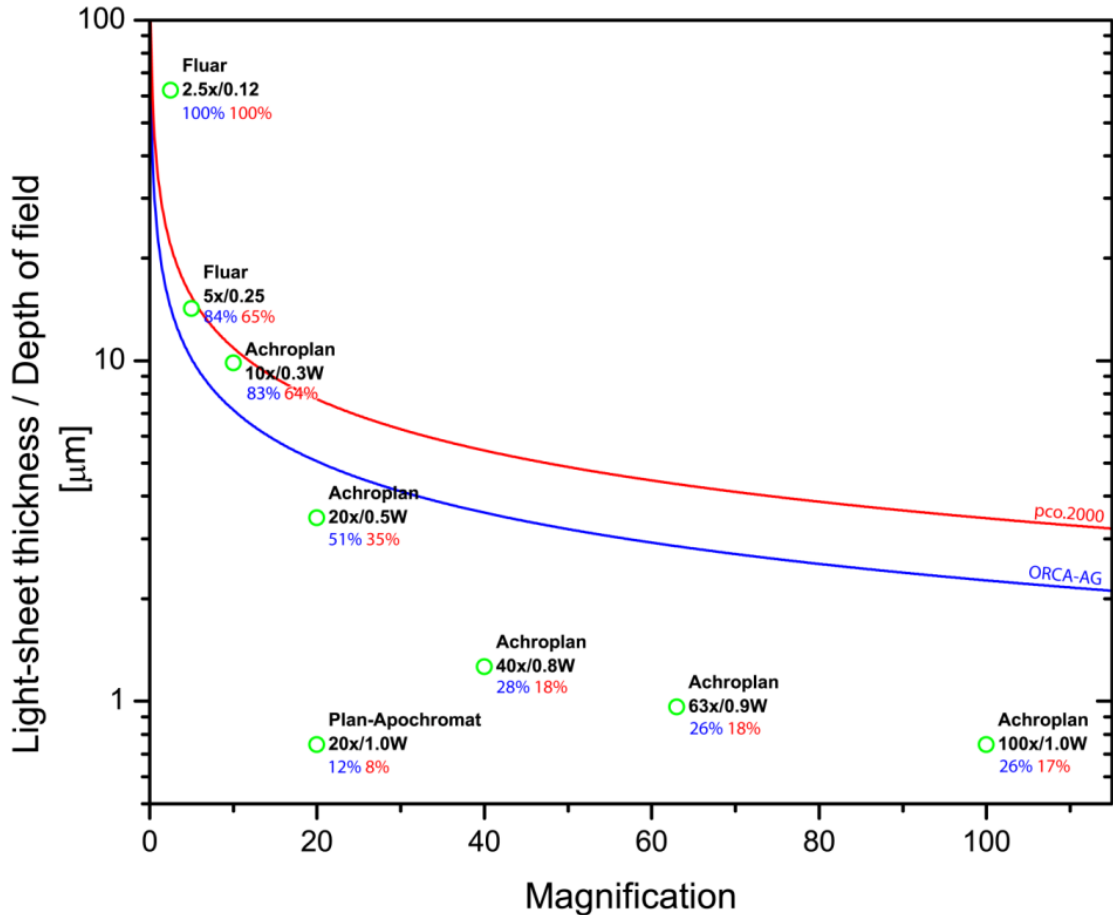


Figure 20: Contrast of the LSFM. The optimal thickness of the light sheet for two most commonly used cameras at EMBL (solid lines) is compared with the depth of focus of common detection objective lens (green circles). For every objective lens, proportion of the illumination power illuminating in-focus volume is calculated for ORCA-AG (blue numbers) and pco.200 (red numbers).

The optimal light-sheet thickness depends on the magnification of the detection objective lens (2.13) and the depth of field, defined by its NA [30,86]. The depth of field of most commonly used objective lenses and the corresponding light sheet thicknesses are shown in the Figure 20. As can be seen from the graph, the light sheet is thicker than the depth of field for objectives with $\text{NA} \geq 0.5$. At $\text{NA} = 1.0$, the proportion of the illumination power in the focal volume drops to a fraction of the total illumination power. However, without planar illumination this ratio would be significantly worse. For example, in case of an evenly illuminated $100\mu\text{m}$ body of cells (see Figure 51), less than 1% of the fluorescence will originate from the focal volume, when collected by a 100x/1.0 objective lens. If planar illumination is applied, this percentage is around 15-30%

(depending on the magnification and the width of the image sensor) independently of the object's size. The percentage is even higher for a detection objective lens with lower NAs.

2.2.3.5 Illumination optics

SPIM (initially called wide-field ϑ -fluorescence-microscope) is EMBL's earliest LSFM implementation. Its first version [71] created the light-sheet by focusing an expanded laser beam by a cylindrical lens (e.g. $f = 150\text{mm}$). The beam is priorly shaped by an aperture to adjust the light-sheet thickness and height (Figure 21a). The cylindrical lenses are poorly corrected for spherical and chromatic aberrations and this method creates a light-sheet of a limited quality.

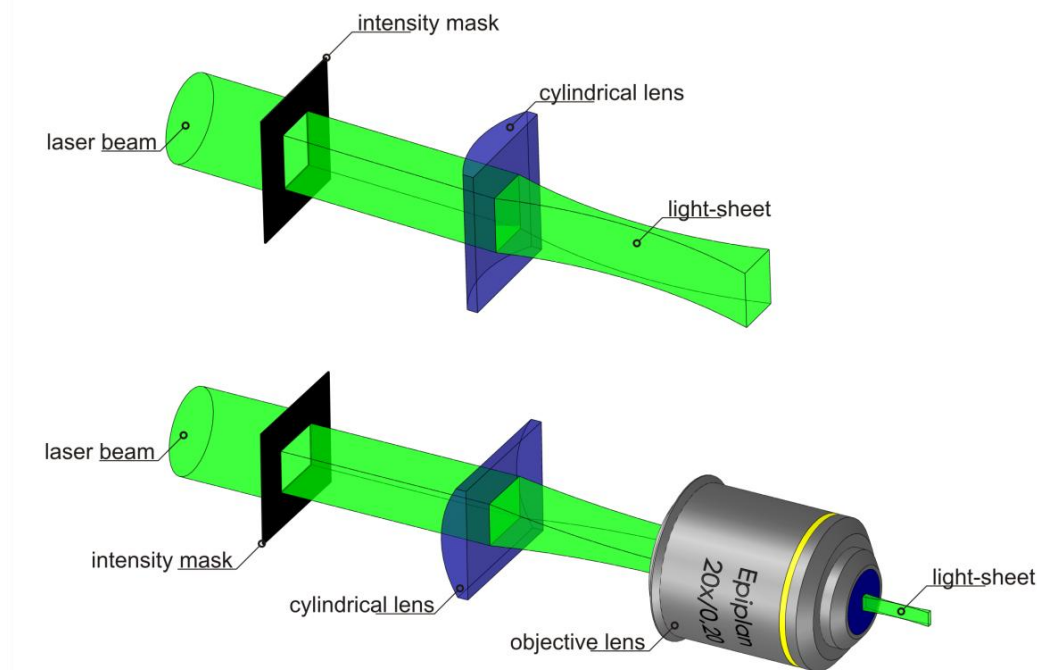


Figure 21: Two types of SPIM illumination. Light-sheet can be created using a cylindrical lens (above) or a cylindrical lens and an objective lens (below). Note different orientation of the cylindrical lens in both arrangements. Intensity mask is used to shape the height and waist diameter of the light sheet.

An alternative method employs an objective lens to focus the light sheet (Figure 21b) [81]. The magnification and NA of the objective lens do not need to be particularly high. The light is normally not focused to a less than $1\mu\text{m}$ thick light-sheet, which corresponds to a NA 0.14 at an illumination wavelength of 488nm. This is why long working distance lens with relatively low NAs (e.g. Carl Zeiss Epiplan 10x/0.20 with a working distance of 17mm) are usually applied, which allow an easy access to the experimental volume. Before the beam is fed into the objective lens' back aperture, it is shaped using a mask and focused by a cylindrical lens to form an intermediate light-sheet at the illumination objective's back aperture. This intermediate light-sheet is oriented perpendicularly to the final light-sheet. The light-sheet is then generated by the objective lens, which is much better corrected than the cylindrical lens. The cylindrical lens is in this setup used to expand the beam along the plane of the light-sheet, which has only a minor impact on the quality of the illumination.

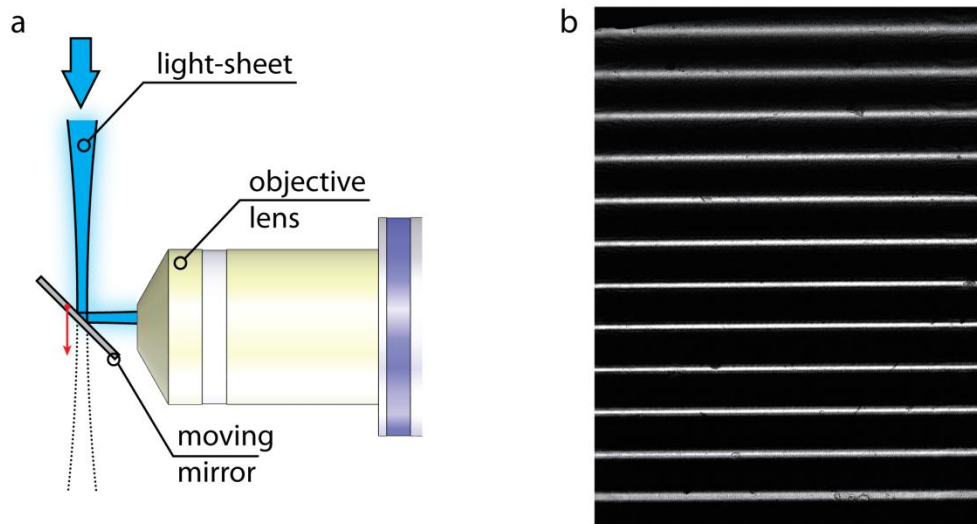


Figure 22: Light-sheet thickness. **a** Properties of the light-sheet are measured using a tilted mirror that reflects it directly into the detection objective lens. By moving the mirror, the light sheet profile at every point in across the field of view can be recorded. **b** A light-sheet profiles at twelve positions spanning the whole field of view (from top to bottom) were recorded and overlaid. The light-sheet is thin in the middle, while it widens and dims at towards the edges. Speckles and dark spots result from dirt particles and scratches on the mirror. The light sheet was created with a SPIM, 488nm argon ion laser, 150mm cylindrical lens and no illumination objective (Figure 21a); recorded by Carl Zeiss Fluor, 5x/0.25 objective lens, OD3 neutral density filter and ORCA EG camera.

Most lasers produce beams with diameters of 1mm or less. Such beams must normally be expanded $1.5\times - 3\times$ before they are shaped with the mask and fed into the objective lens' back aperture, depending on the magnification of the detection objective lens. The adjustment is easiest using a set of exchangeable beam expanders (e.g. Sill Optics S6ASS3105/121, S6ASS3102/121, S6ASS3104/121 for $1.5\times$, $2\times$ and $3\times$ expansion, respectively) or a zoom beam expander (e.g. Sill Optics S6ASS2075/121 for expansion range $1\times - 8\times$).

Earlier versions of SPIM used multiple lasers to provide the wavelengths required for multiple channel (multiple fluorophore) imaging. The beams were coupled into a single beam by set of dichroic mirrors. Newer implementations are based on “white” gas-ion lasers (e.g. Melles Griot 35 IMA 040-220 Ar-Kr-ion laser) to provide most required wavelengths in a single laser tube.

Before the beam is expanded, its spectrum is shaped by an acousto-optical tunable filter (e.g. AA Optoelectronics AOTFnC-VIS). Normally, only one laser line (one wavelength) is transmitted. Its intensity is additionally decreased to achieve optimally illuminated images at given fluorophore concentration and exposure time. Finally, an AOTF is also used to shut down the illumination, when it is not required.

Recently, structured illumination [45,87] was successfully combined with SPIM to additionally improve the image contrast [83]. In SPIM-SI, the modulation mask is inserted into the illumination beam in front of the cylindrical lens. The specimen is now illuminated by a series of stripes, producing a laterally modulated image. The modulation phase is then shifted by a transverse translation of the modulation mask in uniform steps while a series of images (three or more) is recorded. The set is then computationally recombined into a single image. SPIM-SI was shown to produce images with better contrast than SPIM, especially in the case of optically dense specimens [83].

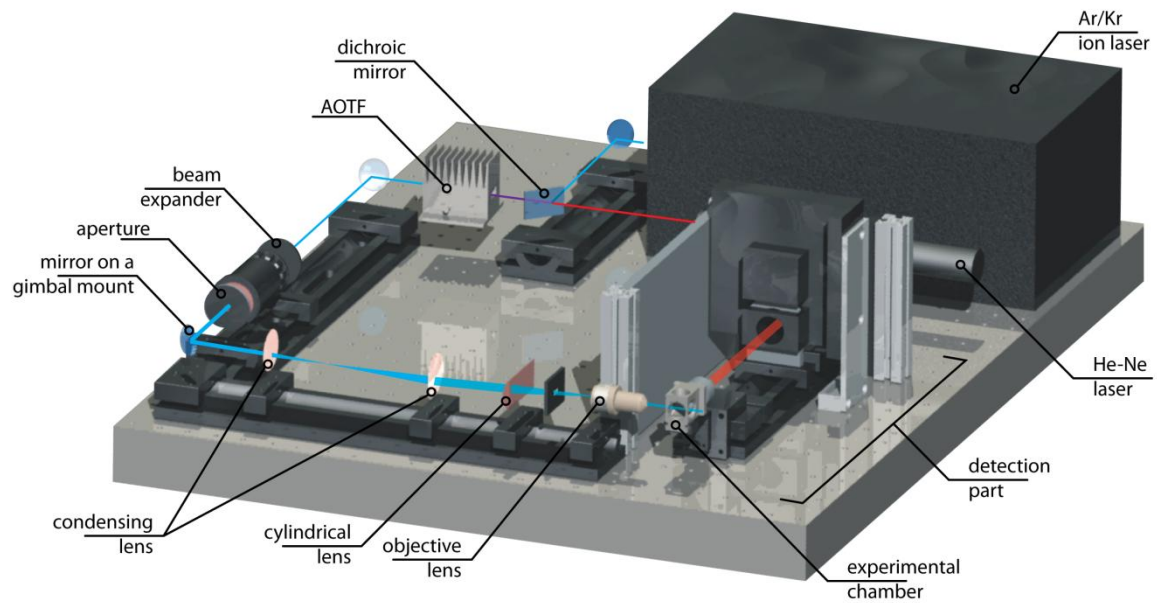


Figure 23: Basic SPIM setup. Beams produced by two lasers are coupled by a dichroic mirror. Only a single wavelength is transmitted through the AOTF with the desired power. Beam is then expanded (beam expander) and shaped to fit the field of view of the detection arm (aperture). Mirror on a gimbal mount and the two condensing lenses are used for parallel translation of the light-sheet (see [81] for details). Light-sheet is formed by combination of a horizontally oriented cylindrical lens and objective lens (as in Figure 21b). Details of the detection part are shown in Figure 13 (camera and tube lens are not visible in this image). Base graphic provided by Christoph Engelbrecht, ETH Zürich.

2.2.4 Digital scanned laser light sheet microscope (DSLIM)

EMBL's more recent implementation of an LSFM is based on a scanning laser beam that creates a dynamic, virtual light sheet (Figure 24) [88]. The illumination system of the DSLM uses no cylindrical lenses so the illumination beam stays circularly symmetric. Instead, a thin section of a fluorescent object is illuminated by scanning a beam across the field of view. Normally, the beam initiates scanning shortly (milliseconds) after the CCD sensor starts collecting light and scans across the field of view once or multiple times during the camera exposure. The image collected on the sensor perceives the integral image, which appears homogeneously illuminated. Apart from the fact, that the DSLM uses a virtual light-sheet to illuminate the specimen, the imaging process is identical to that of a SPIM. Most of the theoretical basis (all but the *Light-sheet height* section in the SPIM discussion above) is therefore similar.

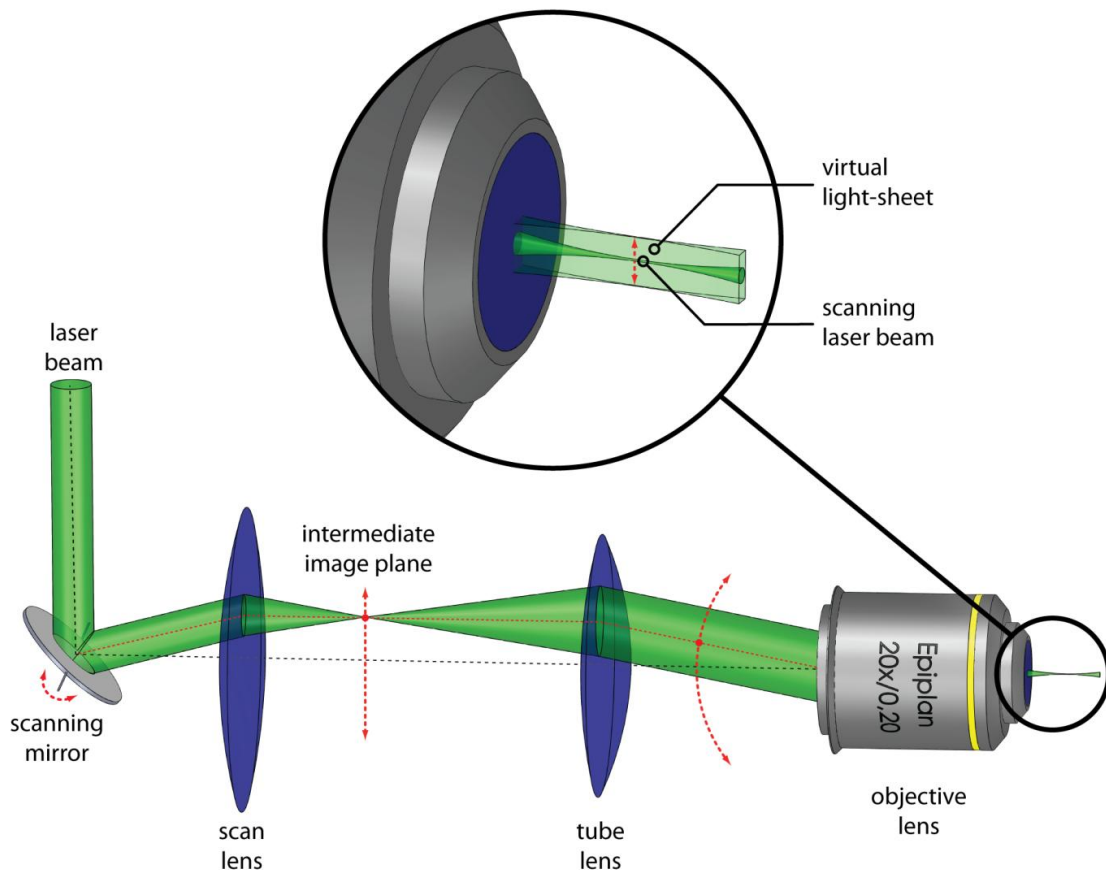


Figure 24: DSLM illumination. DSLM illuminates a specimen by a circularly-symmetric beam that is scanned over the field of view. This creates a virtual light-sheet, which illuminates a section of a specimen just like the SPIM. Light-sheet in DSLM is uniform over the whole field of view and its height can be dynamically altered by changing the beam scan range.

There are two main advantages of such an illumination. Unlike in SPIM, the field of view is now illuminated evenly without wasting any illumination power. This decreases the required exposure time by approx. a factor of 2-3. Additionally, the combination of scanning optics and acousto-optical tunable filter allows patterned illumination. This enables structured planar illumination [83] without a need for a modulation mask.

2.2.5 Specimen translation and rotation

A LSF microscope records an image of a single optical section at a time. The three-dimensional images are created by imaging a set of equidistant sections through the specimen. Optical scanning would involve synchronized translation of the light-sheet and the focal plane of the detection optics to keep them overlapped; mismatch would result in unfocused images. For that reason, the LSFMs built at EMBL keep the optics stationary and translate the specimen through the static light-sheet. Precise translation along the optical axis of the detection lens is realized by the use of a linear translation stage (e.g. Physik Instrumente M-110.2DG with a travel range of 25mm, a minimal step size of 0.2 μ m and a repeatability of 0.5 μ m). Three such stages are normally assembled at right angles; each of them is responsible for movement along one of the Cartesian axes.

A micro rotation stage (e.g. Physik Instrumente M-116DG) is added on top of such a three-dimensional positioner. It allows the rotation of the specimen around an axis, which is essential for multiple-view microscopy (chapter 3). It is critical that the specimen does not deform while it is being imaged along different directions. The axis of rotation is therefore usually vertical, i.e. parallel to the gravity, which nullifies effects of gravity on the specimen during rotation. Such construction also allows the specimen to be immersed into the experimental chamber (Figure 14) from above. An image of a fully assembled translation/rotation stage can be seen in Figure 25.

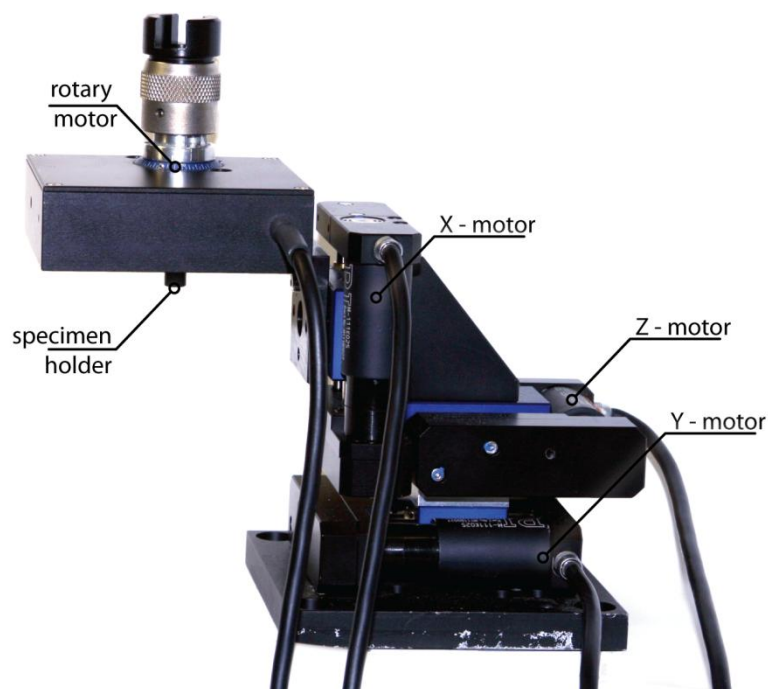


Figure 25: LSFM translation/rotation stage. Stage is assembled from three orthogonally assembled linear translation stages and one rotary stage.

2.3 Specimen preparation and mounting

LSFM seems to be better suited for observation of extended specimens than most other optical sectioning microscopes. Its good penetration depth, long working distance objective lenses without a tradeoff in axial resolution (section 2.2.2) and large experimental chambers (normally approx. 7ml) make LSFM an appropriate tool for the study of biological processes in live macroscopic animals, whole organs and three-dimensional cultured cell-bodies [1]. In the field of cell-culturing, the LSFM fills this niche exactly at the moment, when biologists start to appreciate the importance of performing the experiments in conditions that mimic the environment in living organs [89]. Nothing seems to be farther away from that ideal, than cells grown on a flat glass surface.

During the recent years, completely new ways of preparing specimens for three-dimensional microscopy were established at EMBL. They do away with glass slides and cover slips and rather hold the specimen by means of transparent gels and other transparent non-flat supports.

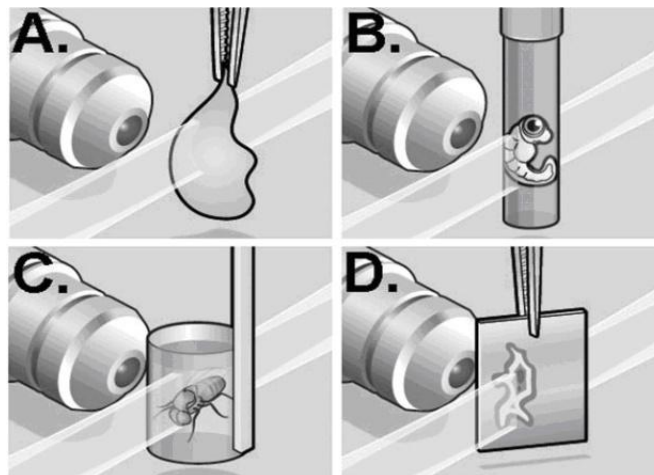


Figure 26: Common ways of specimen mounting in LSFM. Most common ways of mounting a specimen for imaging in a LSFM are **a** clipping, **b** gel embedding, **c** a container and **d** flat mounting

The most common technique is embedding. The specimen is immersed in a liquid gelling agent and sucked into a tube-shaped container (Figure 26b). Once the gelling agent polymerizes, the gel cylinder with embedded specimen is pushed partially out of the capillary and attached to the translation/rotation stage, so that it is immersed in the medium contained in the experimental chamber (Figure 14). There are no glass interfaces between the specimen and the objective lens.

Gels used in LSFM microscopy exhibit low light absorption and refraction indexes very near to that of the water. Most widely used is low concentration (0.5% - 2%), low gelling temperature agarose (e.g. Sigma-Aldrich Type VII agarose, A6560), which is easy to prepare and handle, biologically friendly, sufficiently stiff for most applications and cheap. However, even relatively dilute (1%-2%) agarose gel is not optically completely inert and might cause scattering and astigmatism when objects deep in the agarose cylinder are imaged with high magnification (Figure 27 and Figure 28). On the other hand, low concentration gel (<1%) might provide insufficient stability, especially when very small specimens are being recorded (Figure 29 and Figure 30).

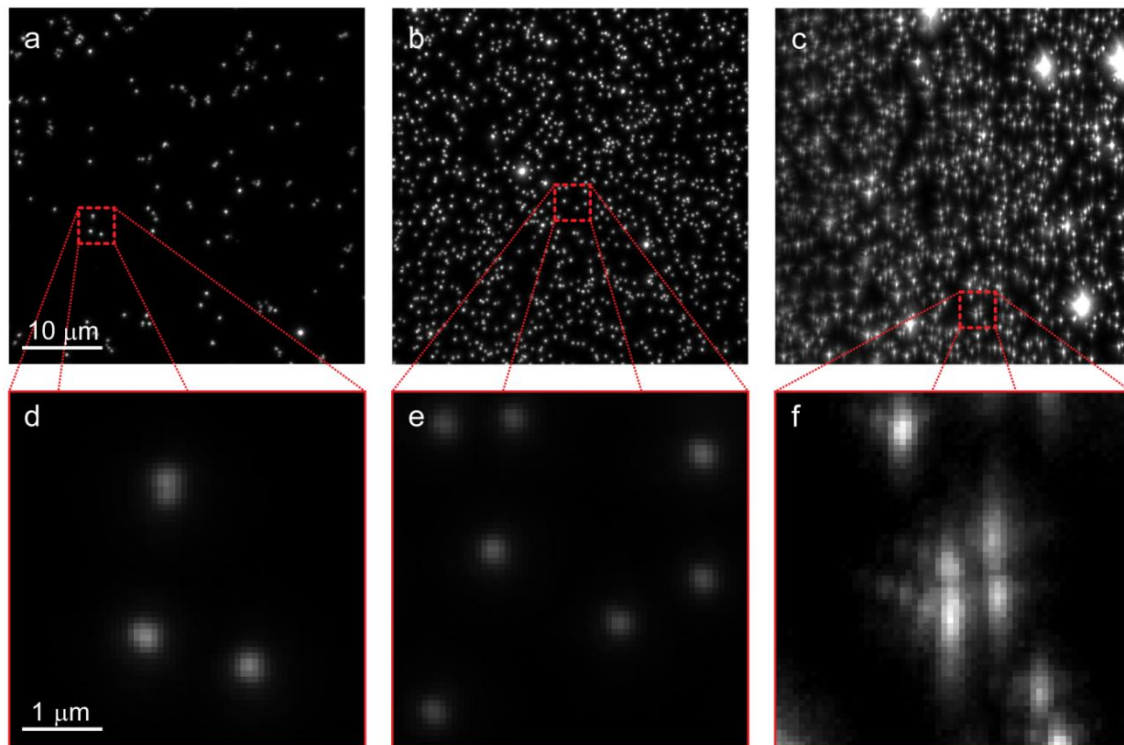


Figure 27: Effect of agarose depth on image quality. Fluorescent beads with a diameter of 93 nm (Invitrogen, FluoSphere Blue) were imaged at surface and 400 μ m deep inside vertically oriented agarose cylinders with two different agarose concentrations. Images **a** and **d**: beads at the surface, 1% agarose concentration; **b** and **e**: beads at surface, 2% agarose concentration; **c** and **f**: beads 400 μ m deep inside 2% agarose block. Images **a**, **b** and **c** show a region of 44 μ m \times 44 μ m, images **d**, **e** and **f** 4.4 μ m \times 4.4 μ m cut from the images directly above them, indicated by the red squares. Imaged with DSLM, Achromat 100x/1.0w objective lens. Illuminated with 378nm UV light, detection between 420nm and 500nm.

Commonly used alternative to agarose is a polysaccharide-based gelling agent Gelrite gellan gum (available from Sigma-Aldrich, G1910), which demonstrates considerably higher light transmission in the visible spectrum than agar (according to the producer). For three-dimensional cell culturing, a set of purposely designed gel-like substrates is usually used. For example gel-like protein mixture, consisting mainly from collagen and laminin (commercially available under trademark Matrigel from BD Biosciences, San Jose, USA) presents the adhering cells with peptide sequences that they are likely to encounter in their natural environments.

Gel-embedding has been extensively used for imaging of *Drosophila melanogaster* (embryo, larvae, pupa and adult) [90], *Anopheles gambiae*, fixed cell cysts and cell aggregates [89], *C. elegans*, *D. rerio* [80,90], *Oryzias latipes*, *Saccharomyces cerevisiae* yeast cells [91] and zooplankton.

Unfortunately, a gel can also restrict physiological movements, such as morphological changes during embryogenesis or cell proliferation, and thus affect the experiment. This can be avoided by containing the specimen in a transparent container, where it is free to move and expand. Such container can be made of gel molded in a special cast or a foil pocket (Figure 26c) made of a polymer that does not disturb imaging. The container can be additionally filled with a more dilute gel that allows physiological movements but would not provide sufficient support if used without the rigid container.

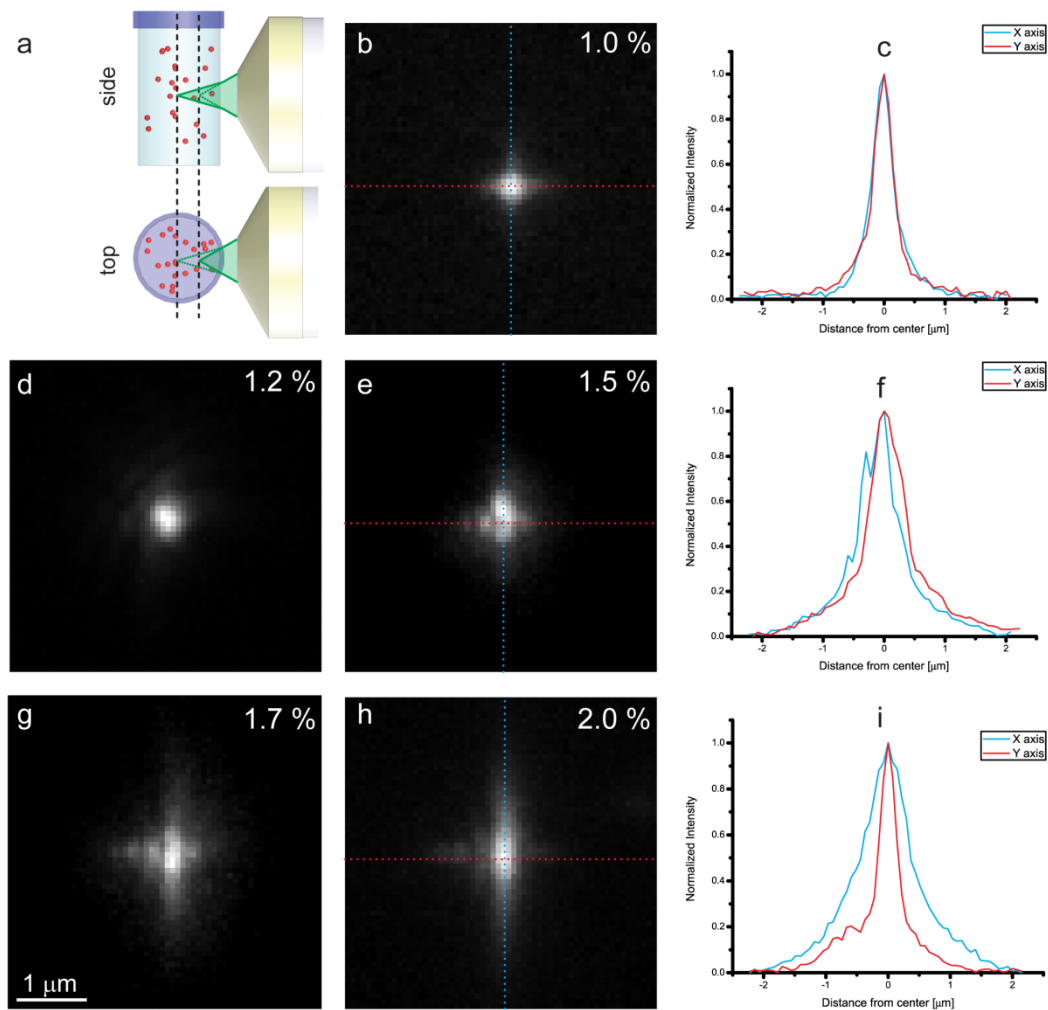


Figure 28: Effect of agarose concentration on image quality. Fluorescent beads with a diameter of 93 nm (Invitrogen, FluoSphere Blue) were imaged 400 μ m deep inside vertically oriented agarose cylinders with different agarose concentrations. Images **b**, **d**, **g**, **e** and **h** show bead images at 1%, 1.2%, 1.5%, 1.7% and 2% agarose concentrations, respectively (as indicated in the upper right corners). Graphs **c**, **f** and **i** show profiles of the images directly to their left (**b**, **e** and **h** respectively) along vertical (X axis) and horizontal (Y axis) directions. Anisotropy in images results from astigmatism caused by lens effect of vertically oriented agarose cylinder with diameter of 0.8mm (diagram **a**) and from horizontally oriented light-sheet, illuminating the beads from the right. Size of the images **b**, **d**, **g**, **e** and **h** is 4.4 μ m \times 4.4 μ m. Imaged with DSLM, Achromat 100x/1.0w objective lens. Illuminated with 378nm UV light, detection between 420nm and 500nm.

This technique has proven to be useful for microscopic imaging of cultured cells growing in three-dimensional extracellular matrices, compression-sensitive specimens (e.g., developing embryos), as well as in vitro microtubule assays [82,92]. The main limitation of this approach is its ability to tailor the chamber size to the specimen. It is difficult to prepare an agarose chambers with an inner diameter of less than 0.5 mm.

The simplest way of mounting a specimen for LSFM is clipping the specimen using tweezers, a clip or a hook made of glass, stainless steel, or plastic (Figure 26a). This is particularly suitable for imaging large specimens such as macroscopic animals and organs (e.g. mouse brain) that are sufficiently rigid to maintain their shape and position. On the other side of the scale, a glass hook proved useful for mounting very small specimens (e.g. yeast cells [91]) dispersed in a transparent

gel. A solid hook provides additional stability to the block of gel, which is critical for high-magnification imaging. However, one has to consider that a solid support can hinder the imaging by interfering with the the illumination or detection light paths.

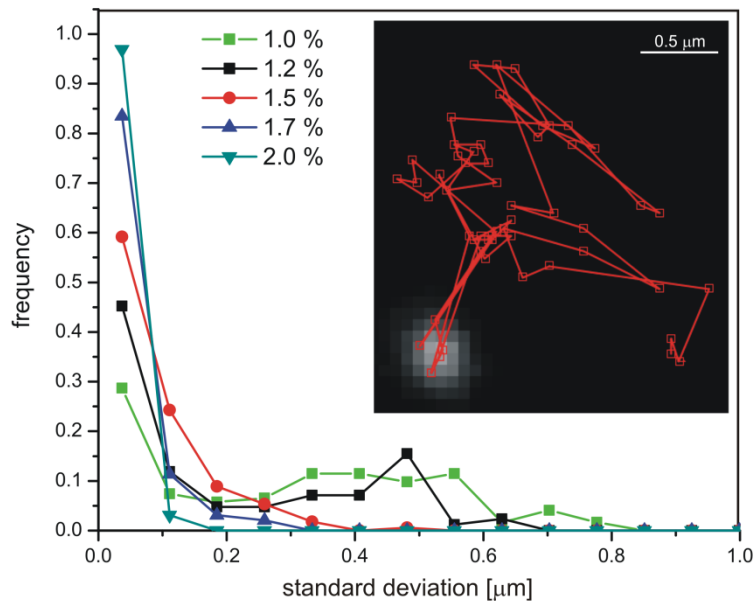


Figure 29: Effect of agarose concentration on bead mobility. Fluorescent beads with a diameter of 93 nm (Invitrogen, FluoSphere Blue) were imaged inside an agarose blocks of different agarose concentrations every second for total of one minute. Beads in the movie were then tracked by a tracking algorithm. The inset shows a typical track for a single mobile bead (1% agarose gel). For each bead, an average positional variance (standard deviation) was calculated. The figure shows the distribution of the beads' variances at different agarose concentrations. At low concentrations, two local maxima are observable: a fraction of immobile beads and a lower and broader peak at around $0.48\mu\text{m}$, which corresponds to diffusion of mobile beads. Knowing beads' size and temperature ($\sim 300\text{K}$) we can estimate effective viscosity felt by the beads to approx. $10\text{-}2\text{Pa}\cdot\text{s}$ or ten times that of water. For details about imaging, see Figure 28.

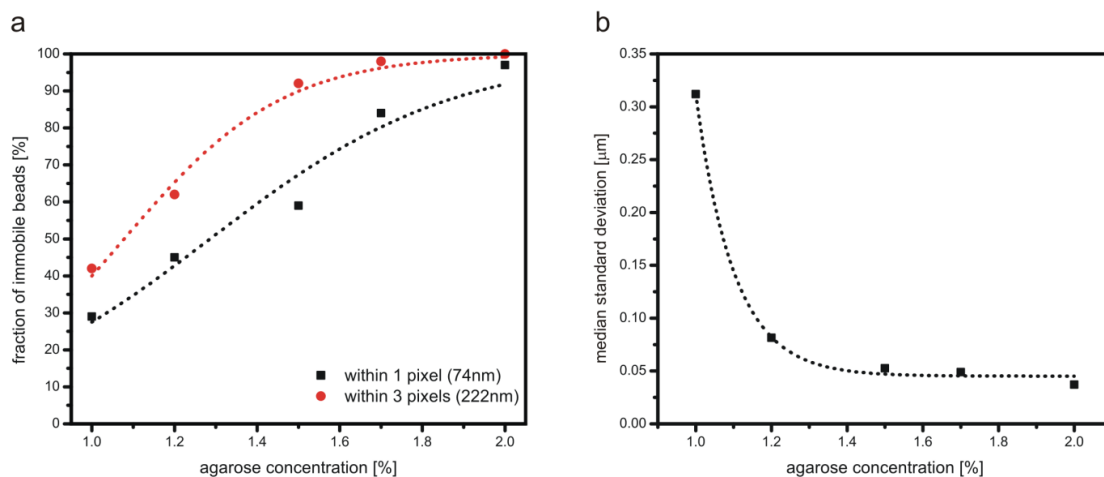


Figure 30: Effect of agarose concentration on bead mobility. The figures show two ways of representing bead mobility at different agarose concentrations, based on the distributions in Figure 29. **a** fraction of immobile beads with two different definitions of immobility. **b** median position deviation at different agarose concentrations. The sharp drop between 1% and 1.2% is probably caused by decrease of a typical pore size in the agarose gel near to the size of the beads.

Specimen can still be mounted on a glass slide that is afterwards clipped in front of the imaging lens (Figure 26d). If the detection axis is horizontal, as is the case in all EMBL's LSFMs, the slide

will be vertically oriented and the specimen must be firmly attached to its surface to prevent it from falling. Furthermore, if a glass slide is oriented normal to the detection lens, it interferes with the light-sheet, especially when very thin specimens are imaged (see Figure 61). In such cases, it is best to rotate the glass slide so that its normal lies somewhere between the illumination and the detection axes.

2.4 LSFM application: imaging of hemocyte migration in a *Drosophila melanogaster* embryo

Cell migration is one of the fundamental processes of life. While bacteria and single celled eukaryotes depend on their movement in search of food and mating partners, cell movement is also essential for development and preservation of multi-cellular organisms. Cell migration is a crucial element of morphogenesis, wound healing and immune response.

In spite of accelerating advances in the field of cell migration, there is still a great uncertainty about the inner workings of cellular motility. Most of our understanding of how cells orient themselves using chemo-attractant concentration gradients [93] and how those gradients are translated into cytoskeleton rearrangements that power cell movement [94,95] comes from experiments with single cells migrating on flat surfaces. While such experiments were a valuable source of insight, their usefulness might be limited by the rather artificial conditions that such migrating cells are subject to. This, together with advances in modern microscopy, has persuaded many biologists that cell migration should be studied where it naturally happens: in the living organisms [96,97].

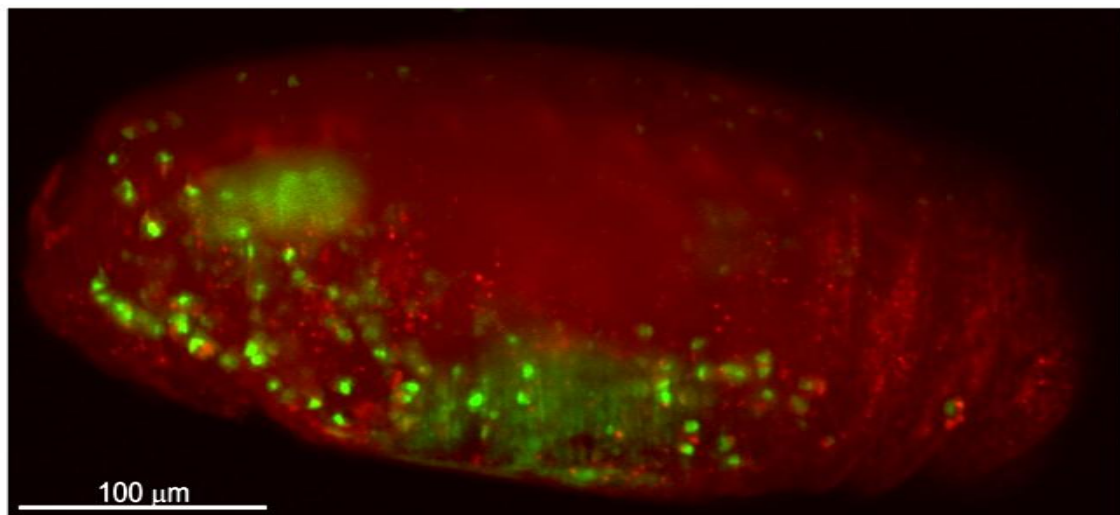


Figure 31: Hemocytes in *Drosophila m.* embryo. Maximum intensity projection through a three-dimensional image of a stage 13 *Drosophila melanogaster* embryo with GFP labeled Hemocyte nuclei (Srp-GAL4, UAS-GFP-nls) in green and cell periphery of all tissues (TubMoeRFP) in red. Images acquired with SPIM, Achromat 10x/0.3w objective lens, excitation at 488nm and 543nm, detection at 500-550nm and above 543nm for GFP and RFP, respectively. The whole three-dimensional image consisted of 670×305×100 volume elements with sampling rates of 0.65μm and 1.3μm in lateral and axial directions, respectively.

2.4.1 *Drosophila m.* hemocytes

There are a number of reasons why *Drosophila melanogaster* seems well suited for cell migration experiments [97]. Probably the single most outstanding reason is the powerful genetic

manipulation developed for *Drosophila*. It allows a quick generation of mutants with perturbed physiological processes and the production of transgenic animals with fluorescently marked proteins [98]. The best studied cell movement processes in *Drosophila* are dorsal closure [99] and the migration of border cells [100], tracheal cells [101], germ cells and hemocytes [102]. Hemocytes are especially interesting as they, apart from following their developmental migration routes, also respond to external impulses like wounding and septic injuries.

The hemocytes are *Drosophila* blood cells involved in the removal of apoptotic and necrotic cells and in response to septic injuries [103]. *Drosophila* hematopoiesis occurs in two waves: in the early stages of embryogenesis and in the larval stages. Embryonic blood cells originate entirely from the head mesoderm of the developing embryo, when they differentiate during stage 8. Around 700 of those cells, called plasmatocytes, are migratory and eventually develop into macrophages while a small fraction, 36 cells, remain stationary and associate with the foregut and give rise to crystal cells.

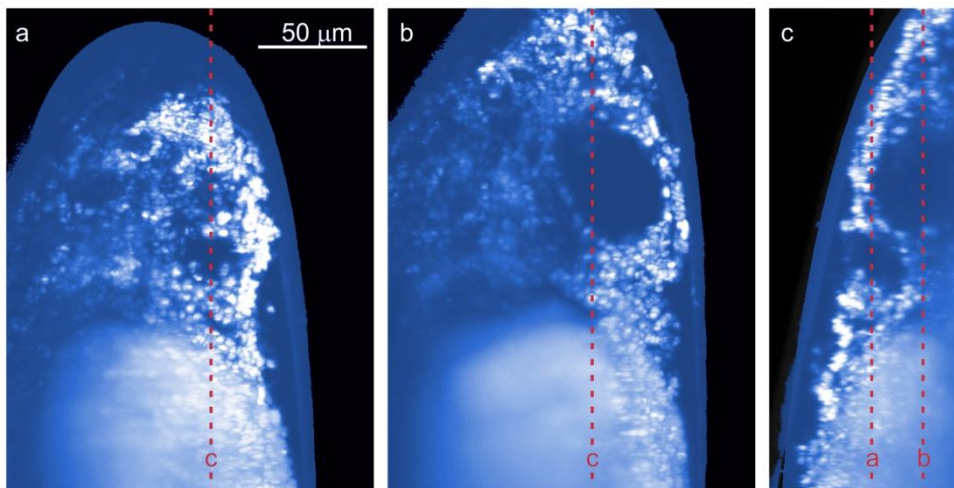


Figure 32: Hemocyte migratory pathways in an intact *Drosophila m.* embryo. Stage 15-16 *Drosophila melanogaster* embryo with GFP labeled Hemocyte nuclei (Srp-GAL4, UAS-GFP-nls) was imaged every minute for a total of 240 minutes, 400 planes per stack. All three-dimensional images of the sequence were then overlaid to reveal Hemocyte migratory routes. Figure shows slices through the embryo at two different depths below the embryo's surface: **a** 34 μm , **b** 53 μm and an orthogonal slice **c**. The blue shadow outlines the boundary of the embryo while red dashed lines show positions of orthogonal slices. Volumes devoid of hemocytes correspond to developing larval nervous system. Images acquired with SPIM, Achromat 40x/0.8w objective lens, excitation at 488nm, detection above 488nm. Three-dimensional image consisted of 1344x1024x400 volume elements with sampling rates of 0.16 μm and 0.65 μm in lateral and axial directions, respectively.

The embryogenetic migration of hemocytes is divided into three phases [104]. In phase I, during embryonic stages 10 and 11, plasmatocytes initiate motility and spread locally throughout the head region. In phase II (stages 12 and 13) plasmatocytes become macrophages and start phagocytosing apoptotic cells, which seems a crucial process for a normal *Drosophila* embryogenesis [105-109]. At the same time they undergo a massive migration and populate the whole body of the embryo. Late in stage 11 they divide into two streams. One of them crosses the amnioserosa to the very posterior end of the germ band, which was brought close to the head during the germ band extension (early stage 11). Those macrophages are then carried with the retracting germ band (late stage 12) to populate the posterior end of the embryo. Both streams meet again in the middle and by migrating dorsally up both sides, spread throughout the whole

embryo. During phase III (stages 15-17) Hemocytes migrate vigorously through the whole body of the embryo, concentrating at the sites of increased cell death where they phagocytose apoptotic cells. Hemocytes in this phase also respond to wounds and phagocytose remains of necrotic cells, bacteria etc.

A GATA homologue transcription factor Srp (serpent) is necessary for the development of both classes of embryonic blood cells [7,8]. By using the Srp promoter we can express any protein of choice selectively in hemocytes. A further specificity can be achieved by using transcription factors Lz (Lozenge) and Gcm (Glial Cell Missing) which are expressed in crystal cells and plasmatocytes, respectively.

Only the embryonic stage of *Drosophila* development is currently well suited for microscopy since it is considerably smaller and more transparent than larval stages or even an adult. The embryo is also intrinsically immobile throughout most of its development. The embryonic tissues are, however, highly scattering, which reduces the penetration depth to only a fraction of the total thickness of the body.

SPIM is well suited for *Drosophila m.* embryo imaging for a number of different reasons. Reduced photo-bleaching allows longer acquisitions at higher frame-rate, which is crucial for imaging of quick cellular processes such as cell migration. Imaging along multiple directions allows us to get more complete view of hemocyte distribution within the Embryo. Examples of hemocyte distributions within *Drosophila m.* embryo as imaged with SPIM can be seen in Figure 31 and Figure 32. Objective lens with 10x magnification were used when the whole embryo was recorded and 40x or 63x for cell-tracking purposes. *Drosophila* embryos are relatively opaque (compared to more transparent animal models, e.g. fish embryos) and good penetration is difficult. SPIM allows good resolution of cell borders up to 20 μ m deep under the surface, while hemocyte nuclei could be resolved even they were 50-100 μ m under the surface (Figure 32).

A UV laser-cutter was also added to a SPIM setup and used for wound-response experiments as proposed in [110,111] (see section 2.4.4).

2.4.2 *Drosophila* transgenes

A hemocyte specific Srp (serpent) promoter [112] was used to drive the expression of fluorescent proteins specifically in hemocytes. It was combined with the Gal4 system [113] to allow a quicker and easier generation of transgenes with hemocyte specific expression.

Initially, a tubulin plus tip binding protein EB1 fused with three copies of GFP was used to visualize migrating hemocytes (Figure 36). Since EB1 is abundant in cell protrusions, the expression of this protein is well suited for study of cell shape changes and protrusion growth during migration, but very cumbersome for cell tracking. Fluorescence is distributed over the whole hemocytes' bodies that rapidly change shape as they move through the surrounding tissue. Furthermore, hemocytes in contact are almost impossible to discern.

Another transgene, featuring nls (nuclear localization sequence) fused with GFP, was therefore constructed. GFP-nls proteins localize to the nucleus, which does not change shape as much as the whole cell and it does not touch the nuclei of the neighboring cells, making automatic segmentation significantly easier.

Each of the lines above was additionally crossed with tubulin promoter driven RFP tagged moesin (TubMoeRFP). This marked the cell periphery of all tissues with RFP, which was used for the visualization of the tissues that hemocytes were migrating through.

Embryos were collected approx. 8 hours after being laid. Their chorion was chemically removed using 50% bleach, as described in [114]. After that, they were embedded in an agarose cylinder [81] and imaged with SPIM.

2.4.3 Automated hemocyte tracking

Most hemocyte migration studies in the past relied on the manual examination of the cells' speed, shape and location in order to reveal possible phenotypes. Manual tracking requires a lot of effort and time, proportional to the number of cells we would like to trace. Manually tracking more than a few dozen cells is hardly feasible. Furthermore, manual tracking is usually only done on 2D images which gives us incorrect measurements when the cells move orthogonally to the plane of the image.

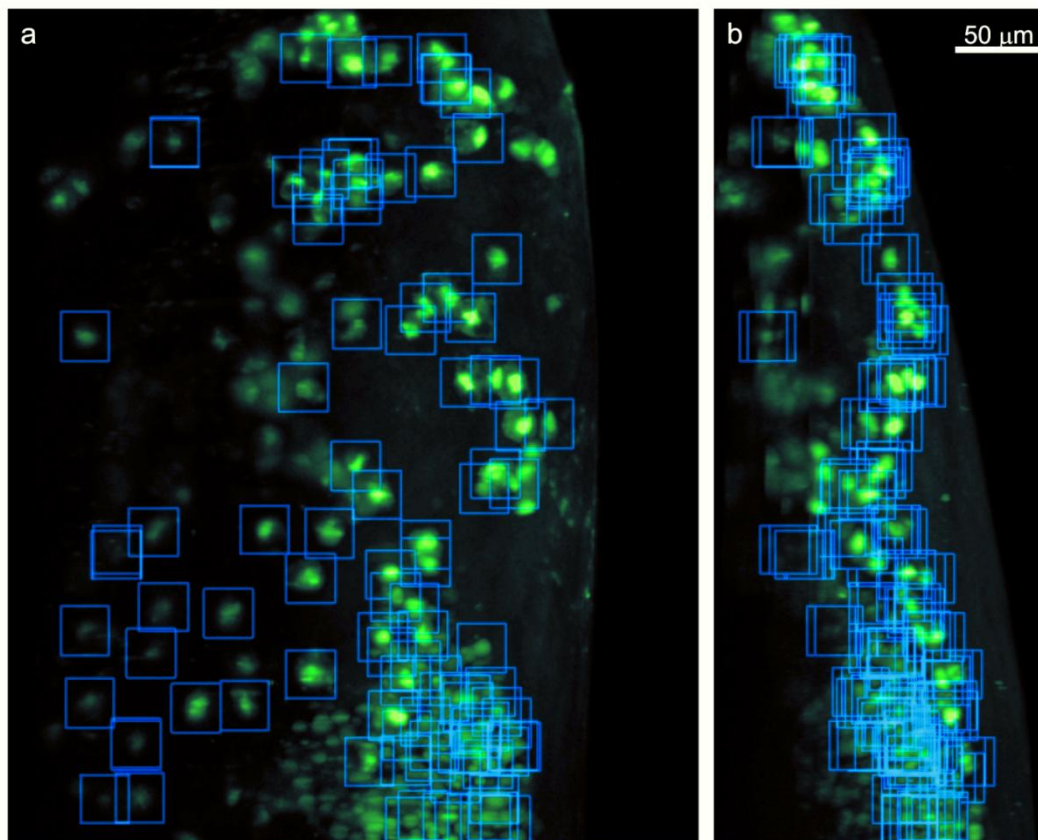


Figure 33: Automated Hemocyte detection in SPIM images. Stage 15-16 *Drosophila melanogaster* embryo with GFP labeled Hemocyte nuclei (Srp-GAL4, UAS-GFP-nls) was imaged every minute for a total of 240 minutes. Automatic Hemocyte detection algorithm was then applied, detecting positions of migrating Hemocytes at every time point (blue cubes). Figure shows top (a) and side (b) maximum intensity projection through an image at a single time point. Images acquired with SPIM, Achromat 40x/0.8w objective lens, excitation at 488nm, detection above 488nm. Three-dimensional image consisted of 1344×1024×400 volume elements with sampling rates of 0.16μm and 0.65μm in lateral and axial directions, respectively.

A more modern alternative is a computer to recognize hemocytes in 3D images of the embryo volume and tracking them while they move. The number of tracked cells can be greatly increased, which gives us better statistics and the possibility to reveal even small changes in the way

hemocytes migrate. After the coordinates of the hemocytes' position are obtained, calculating their speed, direction and other properties of their movement becomes straightforward.

Hemocyte cells were first searched for in every 3D image of a time-lapse. This was done in three steps. A Gaussian filter was first applied to suppress noise. Sub-volumes of the three-dimensional images with sufficient fluorescence signal were then detected in order to identify parts that contain hemocytes and need to be analyzed in the last, more time-consuming part. Finally, watershed based segmentation was applied to each of the sub-volumes to detect and resolve touching hemocytes. Results of hemocyte detection algorithm can be seen in Figure 33.

The algorithm found most hemocyte nuclei close to the surface of the embryo, while its efficiency dropped sharply when hemocytes were more than approx. $70\mu\text{m}$ deep. This could be overcome by analyzing multiple-view imaging (Chapter 3). Furthermore, small auto-fluorescent particles that populate embryonic yolk were usually mistakenly detected as hemocytes.

After the hemocytes were detected in every time-lapse, they were tracked through a sequence of consecutive images. Nuclei in successive images were identified by their positions, fluorescence intensities, sizes and shapes. Unfortunately, the algorithm failed to keep track of the nuclei over prolonged time durations. This was mainly due to two facts, both inherent to hemocytes' nature: i) they tend to move with very relatively high speeds ($1\text{-}2\mu\text{m}/\text{min}$) while abruptly changing direction and ii) they change their shape when squeezing themselves through different tissues. They also often bump into each other or migrate deeper into the body, where they are lost due to image degradation.

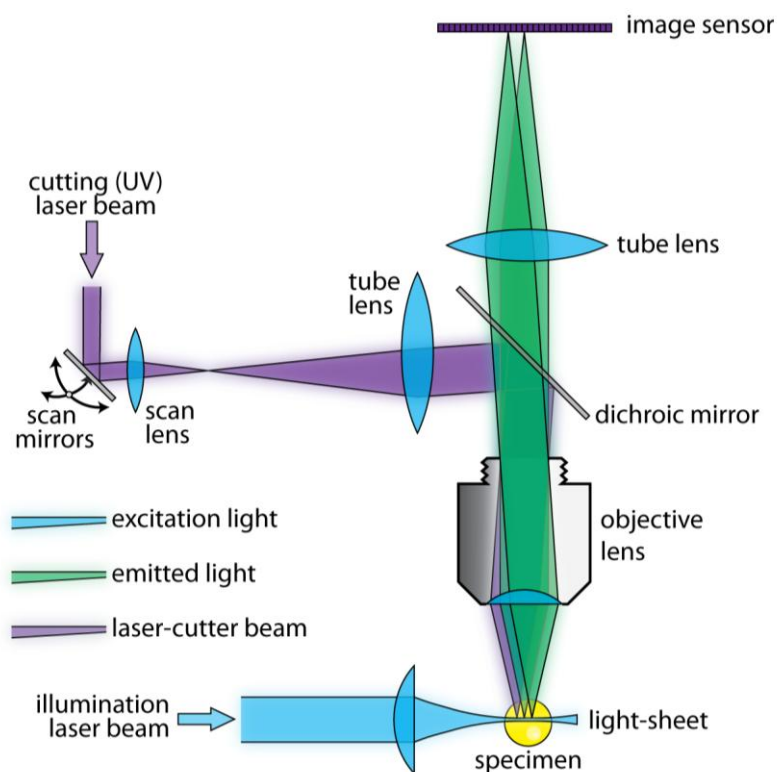


Figure 34: SPIM laser-cutter optical setup. A UV laser is coupled into a conventional SPIM through the imaging objective lens. A dichroic mirror reflecting UV light and transmitting visible light, is inserted in the infinity-space, between the objective lens and tube lens. A scan system consisting of two scan mirrors, f-theta (scan) lens and tube lens is used to steer the beam throughout the field of view.

2.4.4 Laser induced wounding and wound induced hemocyte migration

EMBL's SPIM-based laser-cutter severs cells with a 355nm pulsed UV laser (frequency trippled Nd:YAG) [82]. This particular wavelength seems to be most useful for biological applications. It is short enough to break the chemical bonds in irradiated molecules [115]. Shorter wavelengths are strongly absorbed by DNA and can lead to unwanted damage of genetic material in living samples. The laser is coupled into the detection system of the SPIM just behind the imaging objective lens (Figure 34), which focuses the beam into a spot with diffraction-limited dimensions. Focal volume can be approximated by an ellipsoid with lateral dimensions D_{xy} and axial elongation D_z : [28]:

$$D_{xy} = \frac{2\lambda}{n\sqrt{3 - 2\cos\alpha - \cos 2\alpha}} \quad (2.21)$$

$$D_z = \frac{2\lambda}{n(1 - \cos\alpha)}$$

where α is the angle in the definition of numerical aperture $NA = n \sin \alpha$, λ is wavelength of the light, and n refractive index of the medium. With 40x/0.8W objective lens that were routinely used for experiments, the formulas (1.2) give the following values for the focal volume extensions: $D_{xy} = 0.098 \mu m$ and $D_z = 0.52 \mu m$.

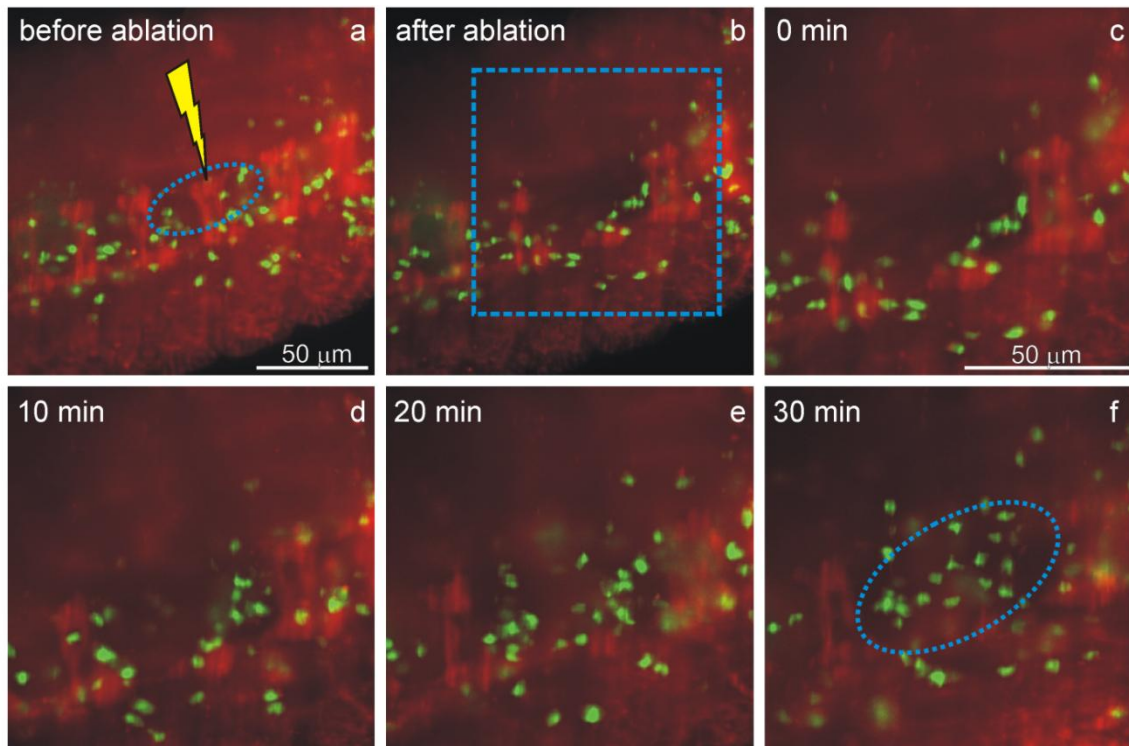


Figure 35: Hemocyte response to a UV laser induced wound. A drosophila embryo with two fluorescent markers (GFP marked Hemocyte nuclei shown in green, RFP marked cell periphery of all tissues shown in red) was used as a model for a wound induced cell migration. A localized internal wound was induced by a series of in total 100 UV-laser pulses with a total energy of $74\mu J \pm 4\mu J$ in the area indicated by the blue ellipse (a). Half of a ventral nerve cord segment was damaged in the process (b). (Embryo was then imaged every minute for the following 30 minutes (c-f). After that time, the irradiated volume (blue ellipse in f) is filled with Hemocytes. Length scale in a applies to b and scale in c applies to d-f. Images acquired with SPIM, Achromat 40x/0.8w objective lens, excitation at 488nm and 543nm, detection at 500-550nm and above 543nm for GFP and RFP, respectively.

The power density in that volume reaches $0.6\text{TW}/\text{cm}^2$ in the peaks of the 470ps long pulses that are produced with repetition rate of 1kHz. Due to nonlinear absorption effects (e.g. plasma formation), a smaller volume around the waist of the focused beam will be strongly perturbed by the irradiation [116]. This volume can be scanned across the whole field of view by a pair of galvanometer-driven scan mirrors.

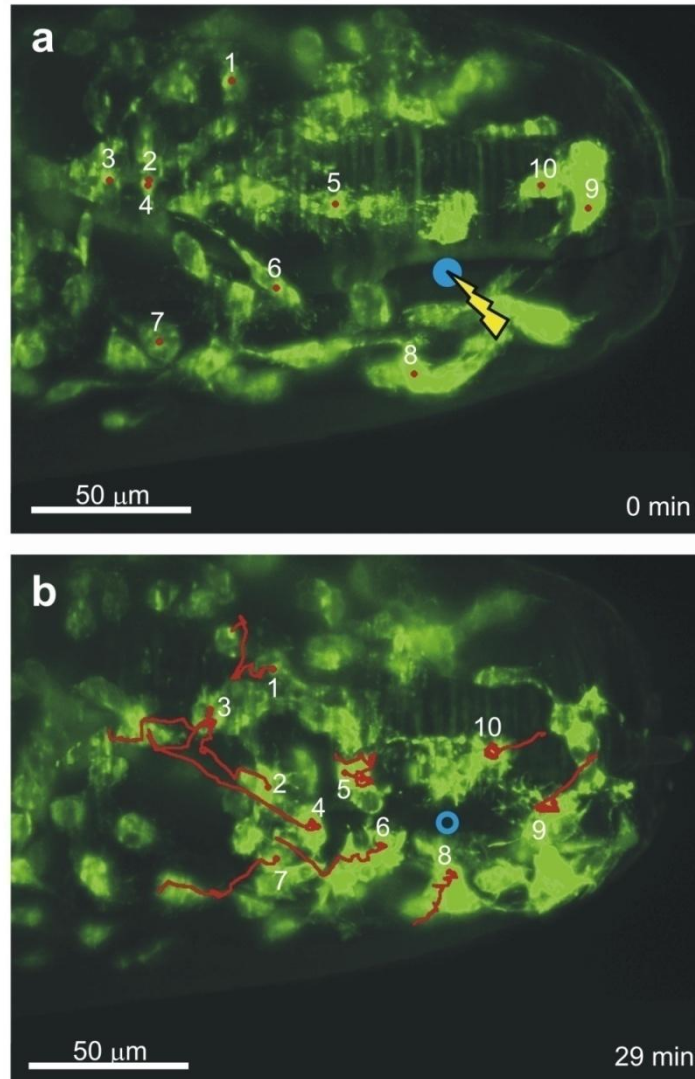


Figure 36: Hemocyte response to a UV laser induced wound. Stage 15 *Drosophila* embryo was irradiated with a UV laser at the location indicated by the blue circle to induce small internal wound (a). GFP marked Hemocytes (Srp-Gal4, UAS-EB1-GFP) were then imaged for 29 minutes. Maximum projections through a stack of fluorescence images before (a) and after (b) laser wounding are shown with migration trajectories of ten randomly selected hemocytes. The data has been recorded with a Zeiss Achroplan 40x/0.8W objective lens, an excitation wavelength of 488 nm and detection above 488 nm. The wound was induced by a series of in total 100 pulses from a frequency tripled Nd:Yag laser (wavelength 355nm) with a total energy of $74\mu\text{J} \pm 4\mu\text{J}$.

The setup allows a tissue perturbation with a very high spatial accuracy. It was shown that by using a similar setup mounted on a wide-field microscope, internal structures of single cells can be selectively destroyed [117] and actin and tubulin filaments cut [118]. When cutting deep inside opaque tissues, the same accuracy can hardly be expected but it is still sufficient to affect only a limited number of cells.

This can be seen in Figure 35, where only about 20 μ m big internal structure deep inside an embryo was injured. Dead irradiated cells have triggered a hemocyte response. Hemocytes have in the course of next 30 minutes migrated to the site of the wound, where they phagocytosed necrotic material. A similar example of wound-induced hemocyte migration is shown in Figure 36.

3 MULTIPLE-VIEW MICROSCOPY WITH LSM



I can conceive few human states more enviable than that of the man to whom, panting in the foul laboratory, or watching for his life under the tropic forest, Isis shall for a moment lift her sacred veil, and show him, once and for ever, the thing he dreamed not of; some law, or even mere hint of a law, explaining one fact; but explaining with it a thousand more, connecting them all with each other and with the mighty whole, till order and meaning shoots through some old Chaos of scattered observations.

Charles Kingsley,
English historian, and novelist
Health and Education (1874)

Observing an object simultaneously from two different viewpoints provides more information, which is why we and many other earthly creatures are equipped with a pair of eyes. In microscopy, multiple-view imaging refers to all imaging processes in which we improve our estimate of the objects three-dimensional density distribution by recording multiple three dimensional images⁸. In the simplest case of a multi-view process we merely take an image N times and improve the signal-to-noise ratio of the measurement by a factor of \sqrt{N} .

A narrower definition of the multiple-view imaging, which we will use in this document, refers to collecting multiple images of the same object along different directions. Such views cover complementary domains in both real and frequency space and therefore offer significantly more information about the imaged object than any of the single views alone.

Different images of a multiple-view set are measurements of the same physical reality - they contain complimentary and partially redundant information about the imaged specimen. In case of fluorescence microscopy, the measured quantity is the spatial fluorophore density distribution. As always, when more than one measurement is available, the best estimate is based on the combined information of the multiple measurements. In the same manner, multiple-view LSM image set is

⁸ *Image* in this text refers to two dimensional and to three dimensional images, also known as *image stacks*. This terminology is a legacy due to the technical implementation of how three dimensional images are acquired or stored on the computer (i.e. as stacks). It is not important for the physical interpretation of the data. A three dimensional image is an evenly sampled three dimensional field and is in this respect analogous to a two dimensional image. Furthermore, changes in the order, in which the voxels (volume elements) of a three dimensional image are saved, do not influence the image. *Image* should be a concept independent of its implementation.

merged into a single estimate of the fluorophore distribution that was likely to have produced the images. The process is called image fusion and is discussed in section 3.5. Resulting estimate of the spatial fluorophore distribution can be again interpreted as an image. It can be used in the same manner and by the same software that is commonly used for visualization and analysis of microscopy images.

In this section, multiple-views microscopy in combination with LSFM is presented. Both essential parts of multiple-view microscopy, image acquisition and digital image fusion, are discussed and original suggestions for overcoming the challenges are proposed. The potential of multiple-view microscopy with LSFM is demonstrated on a number of different specimens, ranging in size from single yeast cells up to a whole *Drosophila Melanogaster* fruit fly (Section 3.6).

3.1 Motivation

There are two fundamental motives for multiple-view imaging in microscopy:

1. Single lens systems such as the conventional epi-fluorescence microscope suffer from the fact that the ratio of lateral and the axial extents of its point spread function is at best $1/3$ (e.g. for $NA = 1.3$) and most commonly somewhere between $1/3$ and $1/15$ (e.g. for $NA = 0.5$, see section 1.6.1 for more)⁹. Thus the frequency space tends to be well filled along the lateral $1/x$ - and $1/y$ -axes but less well filled along the $1/z$ -axis. The observation of an object along multiple directions allows us to fill the frequency space more evenly. Ideally, the multiple views fill the frequency space isotropically. Under such circumstances, the lateral extents of the single view system point spread function dominate the image reconstruction process and the multiple view image fusion provides a better and more isotropic resolution.
2. Images recorded along multiple different directions expose parts of samples that might be obscured along any single direction. Multiple views thus provide us with a more complete representation of opaque specimens. In this case, blurred and dimmed region in one view is replaced by the information from a complimentary view, where this region is sharper and brighter. Final fusion is thus constructed from domains of the single views that offer the least degraded image of that region.

Both motives can be seen as two sides of the same coin: an attempt to improve image's completeness in frequency and real domain. However, image fusion algorithms (section 3.5.4) that achieve each of the two goals differ significantly; no algorithms that do both were reported so far.

Furthermore, in case of an opaque specimen, the two opposing views will optimally complement each other. On the other hand, in case of an ideally transparent specimen, two opposing views

⁹ The divergence between microscope's lateral and axial resolution is less dramatic in LSFM, where the axial resolution at low NAs is governed primarily by the light-sheet thickness rather than by the NA. The ratio between the lateral and axial resolution in LSFM is in the range of 3-4x regardless of the NA (section 2.2.3.3).

produce two identical images. The second image contributes no additional information, since it covers exactly the same disc-shaped domain in the frequency space as the first one. Two views normal to each other contain more information and will produce a fusion with the best and “most” isotropic resolution. With real, semi-transparent specimens, both goals can only be simultaneously achieved, if at least 4 views along two orthogonal axes are acquired.

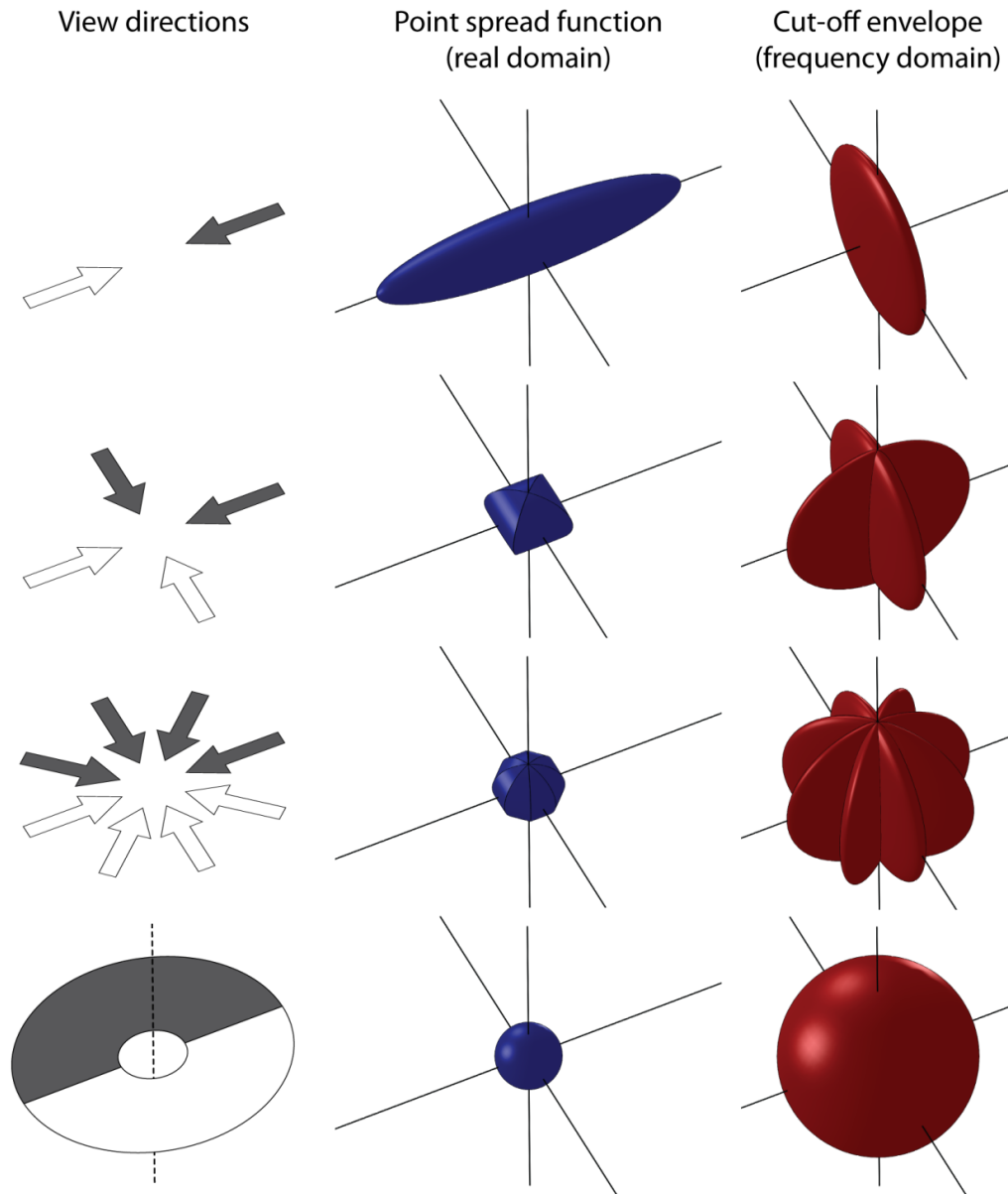


Figure 37: Resolution improvement by multiple-view imaging. First row: Single view (or two opposing views along the same axis) produced by a standard single lens microscope produces PSF that is elongated along the optical axis (blue body). In frequency space, the subspace of recorded frequencies stretches out to higher frequencies in the two lateral dimensions than along the optical axis, forming a disk-shaped cut-off envelope (red body). Second row: If another view is acquired perpendicular to the first one, one more disk-shaped part of the frequency space is recorded, almost doubling the volume of accessible frequencies (red body). PSF is thus reduced, getting a more isotropic shape (blue body). Third row: Recording more views further increases the volume of accessible frequency subspace and improves the shape of effective PSF. Fourth row: in limit of infinite number of views, effective PSF has a spherical shape with isotropic resolution equal to lateral (best) resolution of the single views. Cut-off frequency subspace thus has a spherical shape too, with radius equivalent to lateral radius of single-view frequency envelope disk.

Isotropic resolution

As described in section 1.6.1, the recorded image $I(\vec{r})$ is described as a convolution (*) of the objects fluorophore distribution $O(\vec{r})$ with the microscope's intensity PSF (in case of linear space-invariant image formation):

$$I(\vec{r}) = O(\vec{r}) * PSF(\vec{r}) \quad (3.1)$$

A convolution corresponds to a multiplication of the respective Fourier transforms in frequency space:

$$\mathcal{F}(I(\vec{r})) = \mathcal{F}(O(\vec{r})) \cdot \widetilde{PSF}(\vec{k}) \quad (3.2)$$

If the PSF is symmetric $PSF(\vec{r}) = PSF(-\vec{r})$, then its Fourier transform $\widetilde{PSF}(\vec{k}) = \mathcal{F}(PSF(\vec{r}))$ is a real function. Most true PSFs are very close to symmetric. As a consequence, according to eq. (3.2) the microscope's limited resolution acts primarily as an amplitude filter. The microscope transmits well the low frequencies and increasingly attenuates higher frequencies up to a certain cut-off frequency. Frequencies above the cut-off have no correlation with the object $\mathcal{F}(O(\vec{r}))$, but are determined exclusively by the high-frequency detection noise.

Observation of the same specimen along multiple directions effectively increases the microscope's light-collecting solid angle¹⁰ and therefore increases the inverse-volume of the recorded spatial-frequencies. Just as the $PSF(\vec{r})$ is not spherically symmetric, the same is true for $\widetilde{PSF}(\vec{k})$ – the cut-off frequency is higher in lateral directions than it is axially. In general, the cut-off frequency forms an axially symmetric envelope surface in the frequency space. The surface has a disc-like shape – stretching to higher spatial frequencies along the both lateral axes and being “squeezed” to low frequencies along the optical axis (Figure 37, first row). Adding another view, oriented orthogonally to the first one, almost doubles the inverse volume of the accessible spatial frequencies (second row). Low axial resolution of the first view is complemented by the high lateral resolution of the second view. PSF in the fused image is not elongated anymore, but it is not exactly isotropic either – it still reveals the directions along which the input images were recorded. Adding more views will make it more and more similar to a sphere. In frequency space, this corresponds to making the cut-off envelope more and more sphere-shaped with the final radius corresponding to lateral resolution of the microscope. The whole volume of the sphere will only be filled if an infinite number of views is fused (Figure 37, last row), but image improvement when a number of views above a certain limit is fused, does not justify recording additional images. The ideal number of views depends on ratio between the lateral and axial resolution (i.e. on the NA), durability of our specimen (bleaching, photo-toxicity) and available recording and processing time. Usually, with PSFs characteristic to LSFM, no more than 8 views are required (along 4 axes). Algorithms for resolution-improving image fusion are discussed in section 3.5.4.4.

¹⁰ However, multiple-view fusion is a fundamentally non-coherent image-creation process; input data are intensity images, while phase information is not even recorded. The image formation process is thus fundamentally different from multiple detection lens microscopes that create the image in a coherent manner (e.g. see I⁵M microscope in section 1.6.5)

Completeness of opaque specimen images

Images recorded along multiple directions expose different parts of a specimen. Parts that are obscured behind impervious features look blurred or dimmed or are missed in a single view. Looking along a different direction can reveal those sections, but misses other parts of the specimen. An image fusion created from a set of multiple views in a way that discards the “bad” parts of the images and replaces them with corresponding “good” parts from the alternative views, therefore creates an image that is superior to any of the single views. Such fusion is often referred to as mosaicing and is discussed in section 3.5.4.1.

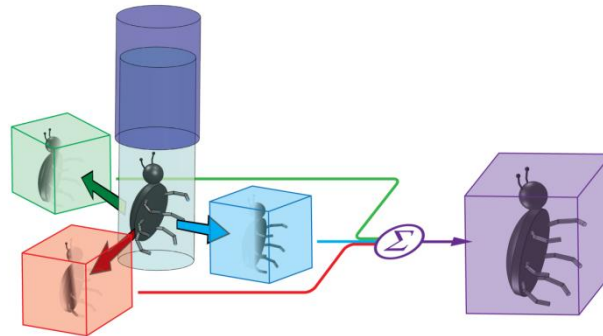


Figure 38: Multiple-view image fusion with an opaque specimen. Each of the three images of a bug reveals a different part of the specimen. After the multiple-view set is fused into a single image, such fusion simultaneously displays all the information contained in single views.

3.2 Multiple-view imaging in microscopy

A stereo-microscope generates two images through a binocular acquired along two tilted axes. It is therefore a two-view microscope. Images perceived by the both eyes are fused in visual cortex of the brain, extracting information about the relative depths from the binocular disparity. The disparity in a stereo microscope is limited by the angular aperture of the objective lens (or the distance between the objectives lenses in a two-lens setup).

In 1974, the specimen rotation in front of single objective lens was proposed as means of resolving complex three-dimensional structures [119]. A purpose-built tilting mechanism was attached to a high magnification (100x/1.25) Nomarski microscope. The specimen was deposited on a glass slide, which limited the mechanical tilting range to $\pm 30^\circ$. However, due to the optical effects produced by the tilted glass interface, the image quality deteriorated at angles above $\pm 15^\circ$. This tilt range was sufficient for a reasonable three-dimensional reconstruction of polytene chromosomes in the salivary gland nuclei of a living adult *Drosophila*. The analysis of the multiple views was performed manually and a chromosome model was constructed from a piece of soft wire.

Multiple-views microscopy gained interest with the advent of computational image processing, the revival of fluorescence microscopy and the application of CCD image sensors in microscopy, which all took place in the late 1980s. The digital image fusion of fluorescence multiple-view images was first reported in 1989 [120]. DAPI stained stage 12 *Drosophila melanogaster* embryos were attached to the surface of a glass capillary. This allowed the rotation over a wider range of angles than if the specimen was lying on a glass slide. In favorable cases, a single *drosophila*

nucleus could be observed as it was rotated from -45° to $+45^{\circ}$. At every angle, a stack of through-focal sections was acquired. Due to bleaching, only two such stacks (normally at 0° and 45°) were obtained before the image quality deteriorated too much. The images were then aligned using a Fourier domain based phase-correlation (section 3.5.3) and fused using a simple frequency domain based algorithm (discussed in section 3.5.4.6).

A number of alternative experimental realizations of multiple-view microscopes were reported since then. They were based on wide-field fluorescence microscopy [121-125], confocal fluorescence microscopy [126,127] and light-sheet based fluorescence microscopy [80,90,128,129]. Imaging a specimen twice along two opposing directions was also demonstrated to compensate the axial attenuation in confocal fluorescence microscopy [130].

Several different methods for specimen mounting that allow the free rotation of the specimen were proposed. They include attaching the specimen to the surface of a glass capillary [120,131] or a thin glass fiber [132], or placing it inside a thin-walled glass or borosilicate capillary [121,126]. Only the second method allows real, unobstructed 2π rotation, but the image quality suffers due to the curved glass-water interfaces between the objective lens and the specimen. This difficulty could be partially mitigated by use of oil-immersion lenses. The specimens for LSFM are normally embedded in a low concentration (up to 1%) agarose gel (see section 2.3), which allows a full 2π rotation and does not harm the image quality.

Some of these implementations employed multiple objective lenses to record multiple views of the same object without having to rotate it. A Double-axis fluorescence microscope [124,125] uses two orthogonally oriented objective lens (20x, NA 0.4). A multiple imaging axis microscope (MIAM) [133] uses four lenses (63x, NA 0.9, water dipping) in a tetrahedral arrangement. Four images acquired through the four lenses were fused by a simple linear algorithm, effectively improving the observation volume by a factor of 3.8 [133].

Computed tomography

In computed tomography (CT), a three-dimensional image is generated by computationally fusing a set of two-dimensional parallel projections along different directions. Fusion is based on the inverse back-projection algorithm (inverse Radon transform) [134]. In analogy to the optical resolution (section 1.6.1.1), a set of multiple-view images with infinitely bad axial resolutions (no axial discrimination) is transformed into a single image with a finite resolution. Thus, computed tomography (CT) can be seen as an extreme version of the multi-view imaging. Since CT methods are based on projection images, they were first used with short wavelength X-ray [135] and gamma-radiation images in the 1970s [136]. CT gave those intrinsically two-dimensional imaging techniques the power to produce three-dimensional images.

Similarly, transmission electron microscope (TEM) generates two-dimensional parallel-projection images of specimen's three-dimensional absorbance distribution. Electron tomography allows three dimensional absorbance map to be reconstructed from a series of two-dimensional TEM images acquired along different directions [137,138]. Low temperatures reduce biological specimens' susceptibility to electron radiation damage, allowing up to $100 \text{ e}^{-}/\text{\AA}^2$. Electron tomography at cryogenic conditions, named cryo-electron tomography [139-141], allows more TEM images to be acquired along more directions (50-150 images covering angles up to $\pm 70^{\circ}$). The resulting three-dimensional resolution is claimed to be in the range of 2-3 nm [141]

Computed tomography was applied to fluorescence microscopy in 2002 and is known as optical projection tomography (OPT) [142]. All these methods generate three-dimensional images from a set of two-dimensional recordings, while LSFM records three-dimensional images by the means of optical sectioning.

3.3 Multiple-view microscopy with LSFM

At least four reasons make LSFM better suited for multiple-view imaging than any other three-dimensional imaging technique:

- specimens are not attached to flat surfaces and can be freely rotated by a full angle of 2π (see section 2.3),
- most LSFM implementations support specimen rotation and translation, which are required for precise rotation and re-positioning during the image acquisition process (see sections 2.2.5 and 3.4),
- the long free working distance (FWD) objective lenses that are usually used in LSFM provide sufficient space for a comfortable rotation of big specimens (e.g. FWD is 2.6mm with a Zeiss Achroplan 10x lens and 1mm with a Zeiss Achroplan 100x, see also section 2.2.1),
- the images produced by LSFM are well suited for further image processing due to their excellent dynamic range and high signal-to-noise ratio.

These advantages were recognized after the first SPIM was built at EMBL Heidelberg and an early fusion technique was already outlined in the initial SPIM publication [80].

Challenges

The amount of data contained in a standard LSFM multiple-view set presents a completely new set of challenges for image processing and multiple-view image fusion. This is due to the high lateral and axial sampling rates (section 2.2.1) that are possible with planar (as opposed to point-scanning) image acquisition, low photo-bleaching rates (Figure 12), improved axial resolution (section 2.2.3), and a high dynamic range (10-12 bit stored as 16 bit images).

For example, a standard image acquired with a Hamamatsu Orca camera (see section 2.2.1) consumes 2.8 megabytes (A PCO camera produces approx. 3x bigger images).

Since axial resolution is 3-5× worse than the lateral resolution, the inter-plane spacing in an image stack is normally selected to be 3-5× bigger than the inter-pixel spacing within every image. The imaged volume normally has a square profile orthogonally to the rotation axis to maximize the intersection of multiple views. Consequently, the number of images in a stack is normally 3-5× less than the number of pixels along single image's edge orthogonal to the rotation axis (e.g. 200-600 planes with Orca EG, up to 800 with PCO.2000). A standard multiple view set consists of 4 or 8 views, up to three fluorescence channels, and a number of time lapses, which totals in considerable amounts of data (see table).

factor	Dataset size (for two different cameras)	
	Orca EG	PCO.2000
Single image	2.8 MB	8.4 MB
Image stack	265× - 512× 700 MB – 1.4 GB	2.2 GB – 7 GB
Multiple channels	1× - 3× 700 MB – 4.2	2.2 GB – 21 GB
Multiple views	4× - 8× 2.8 GB – 33 GB	8 GB – 170 GB
Time-lapse	2× or more from several gigabytes to many terabytes	

The amount of data can be reduced prior to processing by down-sampling the images, by reducing the axial sampling or by reducing the dynamic range of images (e.g. to 8bits), but all this would inevitably mean a waste of useful data and should in general be avoided. The processing of such amounts of data is a demanding task even for up-to-date personal computers. A multiple-view image fusion algorithm is, therefore, required to be memory and processor efficient and provide a robust first solution. A few different approaches to multiple-view fusion of LSFM images were proposed since the initial SPIM publication [90,128,129,143]. They will be outlined in the following two chapters, together with the approaches they suggest for different steps of multiple-view image acquisition and fusion. My contribution will be presented and discussed in greater detail.

3.4 Multiple-view image acquisition

The object must be rotated in order to obtain a multiple-view image set with a single pair of excitation and detection lenses. The axis of rotation must be perpendicular to the optical detection axis. Furthermore, the axis of rotation should be perpendicular to the illumination axis to minimize any effects of eventual obstructions in the illumination pathway (e.g. illumination stripes). The illumination axis, the detection axis and the axis of rotation should, therefore, form a pair-wise perpendicular set of three axes.

The object stability is of utmost importance. If the images that are acquired along different directions do not represent the same quasi-static object, the image fusion will create artifacts or become altogether impossible. The object should, therefore, not change its position and internal structure beyond the anticipate resolution during the time that is required to record a complete multiple-view set of images. This is obviously significant when dynamic live specimens are imaged.

Finally, the specimen should not be deformed during the imaging procedure. Flexible specimens are therefore embedded in a transparent gel that sustains its shape when it is moved and rotated. Deformations of the specimen due to gravity are avoided by rotating the specimen around a vertical axis.

Centering

The field of view (FOV) of an LSFM is in the focal plane of the objective lens and is imaged by the microscope's image sensor via the detection optics. The specimen must be placed inside the FOV in order to be imaged. The specimen is moved through the FOV when recording three-dimensional images.

It is very unlikely that the specimen overlaps with the axis of rotation of the rotary stage. When the rotary stage is rotated, the specimen will therefore move around the axis of rotation¹¹. If the distance between the center of rotation and center of the object is r_0 , then the specimen will travel the distance of $d = r_0\sqrt{2(1 - \cos \alpha)}$ away from its initial position, when the stage is rotated through an angle α . In order to image the specimen after a rotation, the translation motors have to bring it back into the FOV. This can only be done if the position of the FOV and offset between the FOV and the object had been determined. The rotation has to be “calibrated” first. This is normally done using a special calibration specimen that has an easily identifiable point in its structure, e.g. a sharp tip of a pulled glass capillary (see Figure 41). Once the calibration is complete, the calibration specimen is replaced by the specimen of interest.

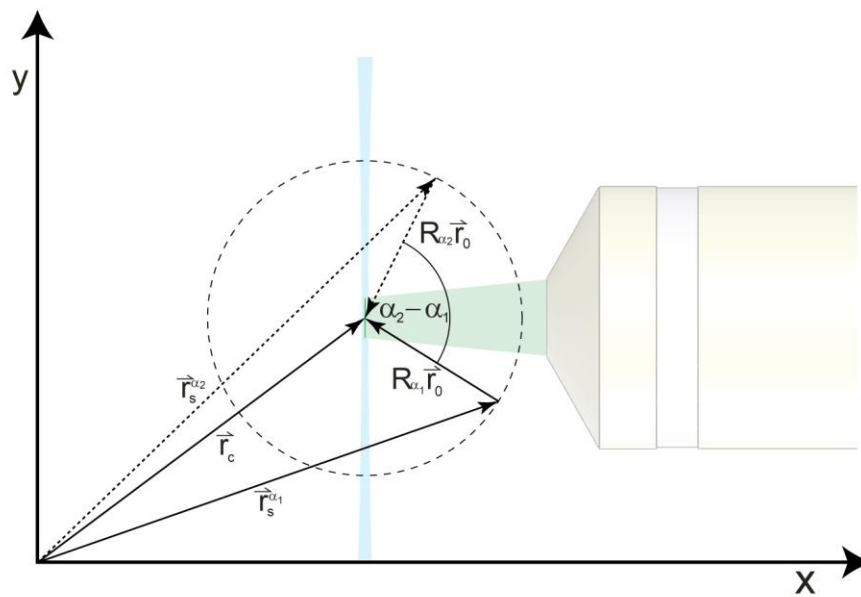


Figure 39: Rotation calibration for LSFM multiple-view imaging. The object of interest is unlikely to coincide with the axis of rotation (\vec{r}_s), i.e. it is displaced for \vec{r}_0 . When rotated, it will therefore undergo a circular motion. The axis of rotation thus has to be moved on a circular path (dashed circle) in order to keep the specimen static in the LSFM’s field of view \vec{r}_c . Solid and dashed vectors show positions of axis of rotation and specimen displacement at two different orientations of rotary motor α_1 and α_2 . The background drawing indicates the position of the light-sheet and the field-of-view of a LSFM microscope.

Let us assume that the axis of rotation is parallel to the \hat{z} axis. In this case the rotation does not affect the Z coordinate so the discussion can be limited to two dimensional x-y case (Figure 39). If \vec{r}_0 is the position of the specimen relative to the center of rotation, α is the angular position of the rotary motor, and \vec{r}_s is the position of the axis of rotation, which can be moved by the stages, the position of the specimen is:

¹¹ This discussion applies to all LSFM implementations, where the rotary motor and consequently the axis of rotation are moved by the translation stages (see section 2.2.5). When the specimen is moved by the translation stages, the axis of rotation is moved by the same distance. The translation stages can thus not be used to move the object into the axis of rotation or vice versa. An alternative approach would be to mount the linear stages and the rotary motor in a different order, i.e. to fix the rotary motor relative to the table and rotate the translation stages. This approach would one to move the specimen relative to the (now static) axis of rotation and eventually allow them to overlap. The repositioning after rotation would thus not be required any more. However in this case, the translation axes would not be aligned with the optical axes of the microscope. Translation along detection axis required to record a stack would require a joint movement of two motors. They would need to be precisely coordinated to produce jiggle-free motion along the detection optical axis. That is most probably why such an approach was (to the best of our knowledge) never implemented yet.

$$\vec{r}_{sp} = \vec{r}_s + R_\alpha \cdot \vec{r}_o \quad (3.3)$$

where R_α is the rotation operator, which in Cartesian coordinates is the rotation matrix:

$$R_\alpha = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \quad (3.4)$$

The specimen is moved into the microscopes FOV (\vec{r}_c) by translating the axis of rotation to:

$$\vec{r}_s^\alpha = \vec{r}_c - R_\alpha \cdot \vec{r}_o \quad (3.5)$$

The required position of the axis of rotation \vec{r}_s^α depends on the angular position of the rotary stage (α) and always lies on a circle with the FOV in the center (\vec{r}_c). To calculate \vec{r}_s^α for a general angle α , \vec{r}_c and \vec{r}_o have to be determined. This is done by moving the same point on the object (e.g. sharp tip of a pulled glass capillary) to the center of the FOV at two different angular positions of the rotary motor α_1 and α_2 . From (3.5) it follows:

$$\begin{aligned} \vec{r}_s^{\alpha_1} &= \vec{r}_c - R_{\alpha_1} \cdot \vec{r}_o \\ \vec{r}_s^{\alpha_2} &= \vec{r}_c - R_{\alpha_2} \cdot \vec{r}_o \end{aligned} \quad (3.6)$$

Solving the system for \vec{r}_c yields:

$$\vec{r}_c = (R'_{\alpha_2} - R'_{\alpha_1})^{-1} \cdot (R'_{\alpha_2} \cdot \vec{r}_s^{\alpha_2} - R'_{\alpha_1} \cdot \vec{r}_s^{\alpha_1}) \quad (3.7)$$

In Cartesian coordinates, after inserting $\vec{r}_s^{\alpha_1} = (x_1, y_1)$, $\vec{r}_s^{\alpha_2} = (x_2, y_2)$ $\Delta\alpha = \alpha_2 - \alpha_1$ equation (3.7) simplifies to:

$$\vec{r}_c = \begin{pmatrix} x_c \\ y_c \end{pmatrix} = -\frac{1}{2} \begin{pmatrix} (x_1 + x_2) + (y_1 - y_2) \cot \frac{\Delta\alpha}{2} \\ (y_1 + y_2) - (x_1 - x_2) \cot \frac{\Delta\alpha}{2} \end{pmatrix} \quad (3.8)$$

A solution exists for every $\Delta\alpha \neq 0$, but the angle difference also determines how an inaccurate measurement of (x_i, y_i) will translate into inaccurate \vec{r}_c . By differentiating (3.8) with respect to x_1 and y_1 , the following relation between the error in \vec{r}_c and the inaccuracy of \vec{r}_s is obtained:

$$E(\Delta\alpha) = \frac{\|\Delta \vec{r}_c\|}{\|\Delta \vec{r}_s\|} = \frac{\sqrt{2}}{2} \left(1 + \left| \cot \frac{\Delta\alpha}{2} \right| \right) \quad (3.9)$$

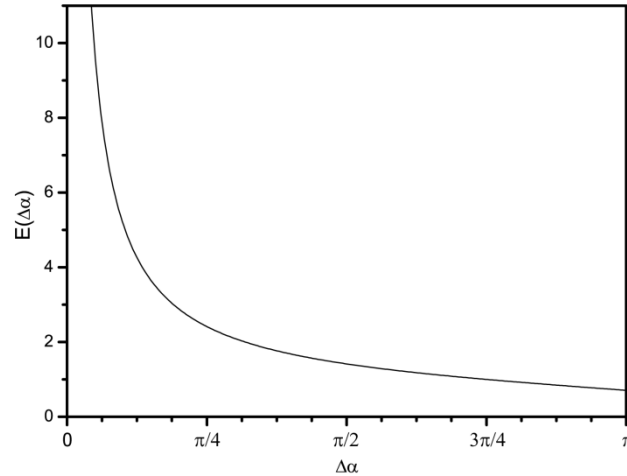


Figure 40: Center of rotation error translation. Calibration error is translated into inaccurate center of rotation depending on the angle difference between the measured views. Errors will affect the accuracy more at small angles ($\Delta\alpha < \frac{\pi}{4}$) and the least if two opposing directions are used ($\Delta\alpha = \pi$). Graph is symmetric over π , error gets bigger again at larger angle differences (not shown).

Figure 40 demonstrates the shape of the error function (3.9). The error will be large at small angle differences ($\Delta\alpha < \frac{\pi}{4}$) while the influence of the measurement inaccuracy will be the smallest when two opposing angular positions of the rotary motor are used, i.e. $\Delta\alpha = \pi$. The equation (3.7) in this case has an especially simple form:

$$\vec{r}_c = \frac{1}{2}(\vec{r}_s^0 + \vec{r}_s^\pi) \quad (3.10)$$

To improve the accuracy further, more than two \vec{r}_s are averaged, with the angles α_i symmetrically distributed around the full angle. For example, early SPIM implementations calculated \vec{r}_c by averaging four measurements at angles $0, \frac{\pi}{2}, \pi$ and $\frac{3\pi}{2}$.

The agarose cylinders most commonly used for mounting small specimens have a diameter of approximately 1mm (see section 2.3). The specimen is normally positioned more or less randomly within the cross-section of the cylinder and will on average lie several hundred microns away from the axis of rotation. The rotation by $\frac{\pi}{4}$ will therefore almost always move the specimen out of the field of view. Finding the object again in each step of the rotation calibration process can be very time consuming.

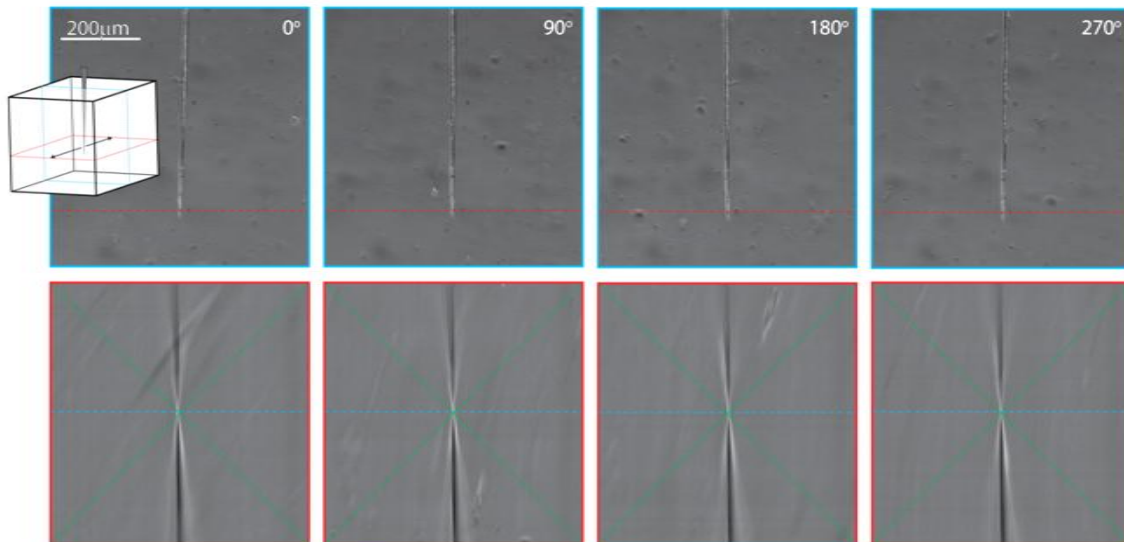


Figure 41: Demonstration of centred rotation. A sharp tip of a pulled glass capillary is imaged in transmission light at four angular positions to probe the accuracy of the rotation calibration. Blue-edged images show lateral cross-sections through the centres of the image stacks and red-edged images show axial cross-sections through image stacks near the end of the tip (see inset diagram – arrows indicate optical axis). Dashed lines show relative positions of the perpendicular cross-sections; green lines indicate the center of the stack. Sharp tip is clearly in focus in the middle of the stack at every angular position.

An alternative approach, implemented in EMBL’s newer LSFM microscopes, executes alignment in an iterative manner. First, a small angle difference (e.g. 5°) is used to get a rough estimate of \vec{r}_c . The angle difference is then gradually increased, using the estimated \vec{r}_c to bring the specimen approximately into the center of the FOV after every rotation. The user is then asked to refine the position, which in turn improves the accuracy of the \vec{r}_c . Normally not more than 5 iterations are required even at highest magnifications (e.g. 100x) to reach the angle difference of π without ever losing the specimen from the field of view.

After \vec{r}_c is determined, the calibration specimen is replaced by the biological specimen. As \vec{r}_c is defined solely by the position of the microscope’s stages relative to the optical setup, changing the specimen does not alter the measured \vec{r}_c . However, the new object of interest is very unlikely to be at the same location as the old one and \vec{r}_0 from eq. (3.5) has to be determined anew. This is done by bringing the new object of interest into the center of the field of view and calculating \vec{r}_0 from eq. (3.5) with the help of \vec{r}_c determined in the calibration process.

Once both \vec{r}_0 and \vec{r}_c are known, the specimen can be rotated by an arbitrary angle around an axis that passes virtually through its center by combination of rotation and translation as defined by (3.5). A demonstration can be seen in the Figure 41. Examples of multiple-view image sets of real biological specimens are shown in Figure 50, Figure 51 and Figure 64.

3.5 Multiple-views image fusion

The most convenient way to harness the potential of the information contained in a multiple-view set is to fuse all views into a single image based on the “good” information from the single views. Such an image can then be analyzed by standard methods used for analysis and visualization of three-dimensional microscopy images. Before the single views are fused, they first have to be registered, i.e. it has to be determined which point in each of the views corresponds of the same

point in the physical space of the specimen (see section 3.5.3). Only after spatial relations between the different single views are established, the information in the single views can be integrated to give a better picture of the specimen.

Understanding how the manner in which the images are recorded in an LSFM provides a good idea about the spatial relation between image spaces of the individual views and the physical space of the specimen. The lateral sampling (pixel pitch) is calculated from the image sensor pixel pitch and the magnification of the LSFM (section 2.2.1), while the plane spacing is defined at the time of image acquisition. Most importantly, each single view was acquired at a well-known angular position of the rotary motor. The single view images will, therefore, look rotated by a defined angle around an approximately (see section 3.5.3.7) known axis, which is approximately parallel to one of the image edges. All this prior information is used to transform all single views into the same coordinate system. This first step of the image fusion procedure is called preprocessing. It produces N images of the same specimen with the same relative orientation and same sampling rates.

A standard image fusion algorithm will, therefore, consist of the following basic steps:

1. **Preprocessing step**, all available a priori information about the geometry of the single views is used to re-orient the images relative to the physical space they represent,
2. **Image registration step**, the information on the spatial relations between individual views is completed by the means of digital image registration and
3. **Final image fusion**, a single three-dimensional image based on the information contained in the whole multiple-view set is produced.

A sketch of the whole procedure is shown in Figure 42. All three steps will be discussed in the following three subchapters. Some processing steps can be combined or omitted. For example, the preprocessing step can be joined with the image registration step. Alternatively, the image registration step could be skipped if the microscope's operation was characterized sufficiently well and geometric relations between the single views are known without having to register the images. However, such an approach is not feasible at this time in the development of LSFM due to difficult technical challenges. In particular, the positioning repeatability and stability of LSFM's stages over long periods of time require some further improvement.

While early image fusion algorithms [80,90,133] were based on Fourier domain (FFT) methods for image registration and final fusion, recently used algorithms were primarily real-space based. There are several reasons why real-domain algorithms are better suited for image fusion algorithms:

- Real space based methods allow space-variant image fusion.
- Real space based algorithms can be easier parallelized and distributed among multiple processors and computers, which will be used to speed-up the processing. Final aim is to fuse the images on-the-fly, as they are recorded.
- Real space algorithms are easier to monitor (intermediate results can be easily examined) and are therefore more robust and easier to optimize.

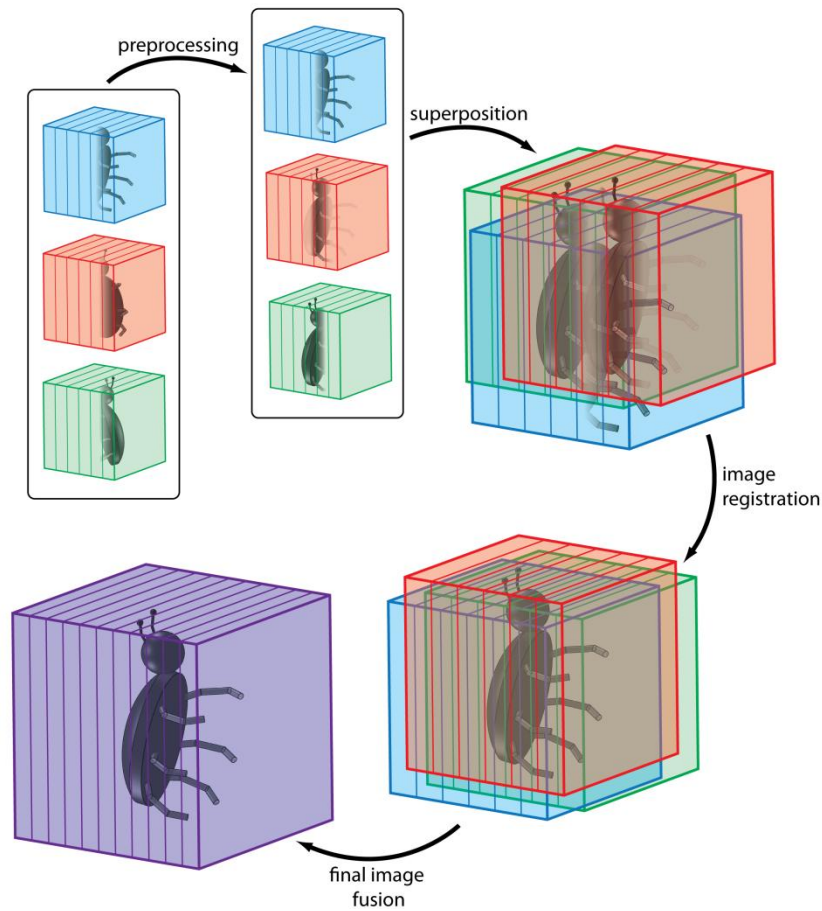


Figure 42: Main parts of image fusion algorithm. **Preprocessing** transforms the images into a common coordinate system, **image registration** aligns the images and finally **image fusion** combines information from individual views into a single image.

3.5.1 Image formation and sampling

Image $I(\vec{r}_i) = I(x_i, y_i, z_i)$ denotes a function that is defined in a discrete set of points \vec{r}_i , a subset of the physical volume that that image represents. The sample points are commonly arranged in a Cartesian mesh of volume elements (voxels), identified by their indices (i, j, k) :

$$I(i, j, k) = I(\vec{r}_0 + i \cdot \vec{i} + j \cdot \vec{j} + k \cdot \vec{k}) \quad (3.11)$$

where \vec{i} , \vec{j} and \vec{k} are mutually perpendicular basis vectors with the lengths corresponding to the voxel pitch distances along the three edges of the image. Vector \vec{r}_0 is the position of the voxel $(0,0,0)$. It is usually chosen such that indexes are non-negative integers: $i, j, k \in 0, 1, 2, \dots$.

The relation between voxel indexes $(i, j, k) = \vec{q}$ and a corresponding physical coordinate $\vec{r} = (x, y, z)$ can however be arbitrary and is generally described by a coordinate mapping \mathcal{M} :

$$\vec{q} = \mathcal{M}(\vec{r}) \quad (3.12)$$

The image values are defined by an object field $O(\vec{r})$ ¹² and by the properties of the imaging apparatus that is used to detect the field. The linear image formation model is most generally described by the following equation:

$$I(\vec{q}) = N \left(\int O(\vec{r}) \cdot h(\vec{r}, \vec{q}) d\vec{r} \right) \quad (3.13)$$

where N denotes a noise process, which in case of LSFM is best described by Poisson statistics. Function $h(\vec{r}, \vec{q})$ ¹³ is called a blurring function and measures how much the field O in point \vec{r} contributes to the image I at point \vec{q} . If a blurring function is space invariant, i.e. it depends only on the difference $\vec{q} - \vec{r}$, it can be replaced with $PSF(\vec{q} - \vec{r}) = h(\vec{r}, \vec{q})$ and (3.13) acquires the familiar form (see section 1.6.1):

$$I(\vec{q}) = N \left(\int O(\vec{r}) \cdot PSF(\vec{q} - \vec{r}) d\vec{r} \right) = N(O(\vec{q}) * PSF(\vec{q})) \quad (3.14)$$

Here, $*$ denotes the convolution operation while $PSF(\vec{r})$ represents a point spread function of the imaging apparatus¹⁴. In the case of multiple views, the same field $O(\vec{r})$ is imaged along n multiple different directions, producing n different images:

$$\begin{aligned} I_1(\vec{q}) &= N(O(\vec{q}) * PSF_1(\vec{q})) \\ I_2(\vec{q}) &= N(O(\vec{q}) * PSF_2(\vec{q})) \\ &\vdots \\ I_n(\vec{q}) &= N(O(\vec{q}) * PSF_n(\vec{q})) \end{aligned} \quad (3.15)$$

Different $PSF_n(\vec{q})$ for $n = 2 \dots N$ are rotated images of $PSF_1(\vec{q})$.

3.5.2 Preprocessing

In this step, a priori information about the position and the orientation of each of the single views $I_i(i, j, k)$ is used. Normally one of the views is selected as a reference and all other images are transformed into its coordinate system. At the same time the images are re-sampled so that they have the same sampling rate in all three dimensions. Additionally, images are rotated so that the axis of the rotary motor is parallel to the \hat{z} axis of the new coordinate system. The last step is important when stacks with different orientations of rotation axes are processed¹⁵. The rest of the algorithm can thus assume rotation axis is parallel to \hat{z} regardless of the orientation of the rotation axis in the input stacks.

Preprocessing is performed independently for each of the views (i.e. without any interaction between the views). It is based on standard image processing techniques, such as scaling and

¹² In fluorescence microscopy, including LSFM, is the object field $O(\vec{r})$ the three-dimensional fluorophore density distribution.

¹³ In this chapter, the blurring function $h(\vec{r}, \vec{q})$ and PSF refer exclusively to their respective *intensity* (as opposed to *amplitude*) kinds.

¹⁴ PSF of an LSFM was calculated and discussed in section 2.2.3.3.

¹⁵ For example, Hamamatsu Orca based LSFMs at EMBL rotate the specimen around \hat{x} axis while PCO.2000 based microscopes rotate it around \hat{y} axis.

rotation, that are described elsewhere (e.g. [144]). Examples of multiple-view sets after preprocessing are shown in Figure 50 and Figure 51.

As mentioned before, the preprocessing step can also be combined with the image registration. This is done by including the a priori information about each of the view's geometry into a starting set of transformation parameters, which is then optimized by the image registration algorithm. This saves a lot processing time and avoids any voxel interpolation, which inevitably blurs the images.

3.5.3 Image registration

As described above, a very precise spatial relation between the single views is required in order to successfully fuse them. A resolution up to $\sqrt{2}$ better than the lateral resolution of the single views (i.e. 3-10× better than axial resolution, see section 2.2.3.3) is expected upon an effective fusion. As a rule of thumb, the single views have to be positioned at least 2x better than the optical resolution of the input images in order to unlock their full potential. In LSFM, this is equivalent to a precision of approx. $0.15\mu m$ with 100x/1.0w objective lens, $0.3\mu m$ with 20x/0.5w and $0.6\mu m$ with 5x/0.25 lens.

The coordinate mappings (3.12) from the physical space of the specimen (\vec{r}) into the coordinate system of the two views (\vec{q}_i) are as follows:

$$\begin{aligned}\vec{q}_1 &= \mathcal{M}_1(\vec{r}) \\ \vec{q}_2 &= \mathcal{M}_2(\vec{r})\end{aligned}\tag{3.16}$$

The mapping \mathcal{M}_i from the physical into the image coordinates cannot be determined easily. What can be determined, however, is a mapping $\mathcal{M}_{1\rightarrow 2}$ between systems of the two images:

$$\vec{q}_2 = \mathcal{M}_2(\mathcal{M}_1^{-1}(\vec{q}_1)) = \mathcal{M}_{1\rightarrow 2}(\vec{q}_1)\tag{3.17}$$

The process of finding the relation $\mathcal{M}_{1\rightarrow 2}$, that optimally maps the first image onto the second image is called image registration or image alignment [145-147].

The registration process is defined by four main components [145]: a) a feature space, b) a search space, c) a search strategy and d) a similarity metric. *Feature space* defines which information from two images will be used to evaluate their match. Our approaches were based mainly on area-based methods, i.e. comparing intensity levels in both images over corresponding extended regions, though registration of segmented images (feature based method) was investigated as well (section 3.5.3.7).

Search space is the class of transformations $\mathcal{M}_{1\rightarrow 2}$ considered for registration of pairs of images. Transformations are described by a coordinate mapping, defined by a limited set of parameters. For example, translation in 3D is described by 3 parameters and affine transformation by 12 parameters.

The *search strategy* is a method for finding the set of parameters that produces optimal fit between the two images. It is an optimization problem of maximizing the similarity metric in the minimal number of metric evaluations. We have used primarily a *regular step gradient descent* optimization, where the space of transform parameters is explored in steps of uniform length in

the direction of the steepest gradient. When a locally best minimum is found, the step size is decreased and the procedure is repeated. Process is interrupted when the step size becomes smaller than the required precision.

The image registration algorithms normally align only two images at a time. Multiple image registration is constructed from multiple pair-wise registrations. Early approaches [90,133] registered every image with an intermediate image, constructed from all the other images of a multiple-view set. The process was cyclically repeated, until additional repetitions brought no further change. This scheme gave good results but it was time demanding, since every image had to be registered with the rest of the set multiple times.

The new approach is based on the observation, that neighboring views (i.e. views along directions intersecting at a sharp angle) contain the most of the common information, which is required for a registration. Multiple-view images are thus grouped into pairs of images that were recorded along neighboring directions (Figure 43). Two images of each pair are then aligned with each other and the resulting parameters are stored. The pair is then fused into an intermediate image. The number of intermediate images is, therefore, half of the original input images. The procedure is then repeated with the intermediate images. The number of images being registered is halved during each iteration. Process is interrupted once there is only one image left, containing all initial input images. Transformations, that each input image had to undergo to fit into the final image, are the result of the registration and are used for the final fusion.

The fusion algorithm used to obtain the intermediate images can be very simple and quick, i.e. an arithmetic mean of the aligned input images (section 3.5.4.10). Alternatively, an image constructed from the brightest pixels of input images (section 3.5.4.2) is a better choice with very opaque specimens. This method is most efficient if the number of the views N in a set is a power of two ($N = 2^x$) in which case it will require only $N - 1$ pair-wise registrations.

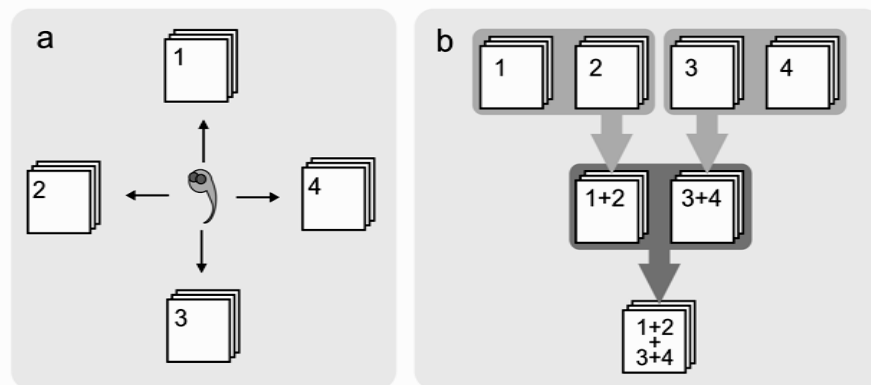


Figure 43: An example of the pair-wise alignment of multiple-views. **a** Four images of a fish are recorded along four perpendicular directions. **b** Images are aligned in pairs of two neighbours and then fused into intermediate stacks. This stacks are then aligned and fused again, until there is just one stack left. While four views are aligned in this example, we can use the same procedure to align any number of them.

Similarity metric

Finally, the similarity metric measures how well two images fit, after one of them has undergone a geometric transformation. Similarity metric is a function of two images and usually produces a scalar number. In the remainder of the chapter it is assumed, that one of the images being fed to

the metric (e.g. I_1) has already been transformed. Optimization of the transform parameters with regard to the similarity metric is expected to improve the fit between the images.

We have used primarily the cross-correlation and normalized cross-correlation [148,149], and mutual information [150,151] metrics. The latter is primarily used for registration of multi-modal images, but it works also well with images with same modality.

3.5.3.1 Cross-correlation

The cross-correlation based measure of the degree of similarity between two images (I_1 and I_2) is motivated by the Euclidian distance between the two images' intensity levels:

$$\begin{aligned} D^2(I_1, I_2) &= \sum_{i,j,k} (I_2(i, j, k) - I_1(i, j, k))^2 \\ &= \sum_{i,j,k} I_1^2(i, j, k) + \sum_{i,j,k} I_2^2(i, j, k) - 2 \sum_{i,j,k} I_1(i, j, k) I_2(i, j, k) \end{aligned} \quad (3.18)$$

The first two terms of the expression depend only on individual images and do not measure similarity. This is done by the final, cross-correlation term:

$$C(I_1, I_2) = \sum_{i,j,k} I_1(i, j, k) I_2(i, j, k) \quad (3.19)$$

This simple *cross-correlation* metric has a number of problems: i) the range of C depends on the images' size, ii) the expression is not invariant to scaling and shifting one of the images' values (i.e. $I' = aI + b$) and iii) the cross-correlation C grades similarity between an arbitrary pattern and any high intensity area (regardless of its structure) always high. Cross-correlation might therefore drive the image registration in direction of high overlap between bright areas instead of a good match between the images' features. These shortcomings can be overcome by normalizing images' ranges to unit length:

$$NCC(I_1, I_2) = \frac{\sum[(I_1 - \bar{I}_1)(I_2 - \bar{I}_2)]}{\sqrt{\sum(I_1 - \bar{I}_1)^2 \cdot \sum(I_2 - \bar{I}_2)^2}} \quad (3.20)$$

where the operations inside the sums are performed on a pixel-per-pixel basis and \bar{I}_n represents the average intensity of image I_n . The *Normalized cross-correlation* NCC eq. (3.20) produces value 1 only in case of a perfect match between both images (i.e. $I_2 = aI_1 + b$ where $a > 0$) and stays in the range $[0,1]$ otherwise (see [148] for a mathematical derivation and proof).

3.5.3.2 Phase correlation

If two images are only translated relative to each other (i.e. $I_2(\vec{q}) = I_1(\vec{q} + \vec{q}_t)$) then the displacement, which provides the globally best cross-correlation (3.19) value, can be efficiently calculated using the *phase correlation* registration. Due to shift theorem, Fourier transforms $F_n = \mathcal{F}(I_n)$ of two translated images are connected by the following equation:

$$F_2(\vec{\omega}) = F_1(\vec{\omega}) \cdot e^{-i\vec{\omega} \cdot \vec{q}_t} \quad (3.21)$$

if $I_2(\vec{q}) = I_1(\vec{q} + \vec{q}_t)$. The phase difference is easily calculated from the images' cross-power spectrum:

$$\chi(\vec{\omega}) = \frac{F_1(\vec{\omega})F_2^*(\vec{\omega})}{|F_1(\vec{\omega})F_2^*(\vec{\omega})|} = e^{i\vec{\omega}\cdot\vec{q}_t} \quad (3.22)$$

The inverse transformed spectrum has a peak at a position corresponding to $-\vec{q}_t$ and is zero otherwise:

$$\mathcal{F}^{-1}[\chi(\vec{\omega})](\vec{q}) = (2\pi)^3 \delta(\vec{q} + \vec{q}_t) \quad (3.23)$$

where δ stands for Dirac delta function. If the images do not match exactly, the cross-correlation function (3.19) will be have the global maximum at \vec{q} , where $\mathcal{F}^{-1}[\chi(\vec{\omega})](\vec{q})$ from (3.23) has the maximum value.

The strength of the phase correlation method is that it finds the global solution for transformation optimization problem in a constant time. The fourier transforms are efficiently calculated using fast Fourier transformation (FFT, see e.g. [152]). The time needed for a phase correlation registration can, however, be longer than the time required by iterative optimization scheme if the iterative optimization is started from a point reasonably close to the global minimum. The phase correlation method was further expanded to include the registration of images that are related to more complex transformations than only translation (e.g. translation and rotation [153]), but those were not tested with LSFM images yet.

3.5.3.3 Registration by maximization of mutual information

Cross-correlation based metrics rely on the assumption that the intensities in interrelated images are linearly correlated. This is true with single fluorescence channel LSFM images, while the assumption is not true in case of multimodal images, e.g. images recorded via different fluorescence channels (i.e. different fluorophores). The image registration by maximization of mutual information [150,151,154] is a relatively new approach to the problem of alignment of multimodal images, based on information theory [155].

The voxel intensities in the two registered images can be regarded as two random variables I_1 and I_2 with marginal probability distributions $p_{I_1}(i_1)$ and $p_{I_2}(i_2)$. The measure of similarity between the two images is summed in a joint probability distribution $p_{I_1 I_2}(i_1, i_2)$, which is a probability that a voxel at a randomly chosen location in I_1 will have the intensity i_1 , while at that same location in I_2 , the intensity will be i_2 . No correlation between the two images means, that the distributions defined by I_1 and I_2 are statistically independent and the joint probability distribution is a product of their independent marginal distributions $p_{I_1 I_2}(i_1, i_2) = p_{I_1}(i_1) \cdot p_{I_2}(i_2)$. On the other hand, if there is a uniform relation between both intensities $i_2 = T(i_1)$ then $p_{I_1 I_2}(i_1, i_2) = p_{I_1 I_2}(i_1, T(i_1))$ and the images are said to be maximally dependent. In case of statistically independent image intensities, the intensity of a pixel in I_1 bears no information about the intensity of a pixel at the same location in I_2 . Inversely, in case of maximally dependent images, the information about the pixel intensity in image I_1 fully determines the intensity at the same location in I_2 .

Normally, the similarity between two images lies somewhere between the two extremes described above. Mutual information based similarity metric estimates how statistically related both images are by measuring the distance between the measured joint distribution $p_{I_1 I_2}(i_1, i_2)$ and the joint distribution the two distributions $p_{I_1}(i_1)$ and $p_{I_2}(i_2)$. would present if they were statistically independent (i.e. $p_{I_1}(i_1) \cdot p_{I_2}(i_2)$). This is done by the means of the Kullback-Leibler measure [156], i.e.:

$$MI(I_1, I_2) = \sum_{a,b} p_{I_1 I_2}(i_1, i_2) \log \frac{p_{I_1 I_2}(i_1, i_2)}{p_{I_1}(i_1) \cdot p_{I_2}(i_2)} \quad (3.24)$$

The mutual information MI is zero in the case of statistically independent images and positive otherwise (it is bounded by the entropies of the individual images [151]).

3.5.3.4 Histogram implementation of area-based similarity metric optimization

Cross-correlation (3.19) , normalized cross-correlation (3.20) and mutual information (3.24) metrics are determined exclusively by the joint image intensities, i.e. intensities of the analogous pixels. It is not important, how the intensities are distributed in the image, the only thing that matters is how does an intensity at any location in image I_1 relates to the intensity at the same location in image I_2 . This relation between intensities of two images can be condensed into a joint histogram $\mathcal{H}_{I_1, I_2}(a, b)$, which measures the number of occurrences of intensities a and b at the same location of I_1 and I_2 , respectively. The similarity information from two LSFM images that take 0.5 - 2 gigabytes each, is thus compressed into a two dimensional histogram with a size between 256 kilobytes (8 bit representation – 256x256 histogram) and 256 megabytes (12 bit – 4096x4096 histogram).

The joint and marginal probability distributions required to calculate mutual information (3.24) can be readily obtained from such a histogram:

$$\begin{aligned} p_{I_1 I_2}(i_1, i_2) &= \frac{\mathcal{H}_{I_1, I_2}(i_1, i_2)}{\sum_{i_1, i_2} \mathcal{H}_{I_1, I_2}(i_1, i_2)} \\ p_{I_1}(i_1) &= \sum_{i_2} p_{I_1 I_2}(i_1, i_2) \\ p_{I_2}(i_2) &= \sum_{i_1} p_{I_1 I_2}(i_1, i_2) \end{aligned} \quad (3.25)$$

Similarly, the normalized cross-correlation (3.20) can be calculated directly from the joint histogram:

$$NCC(I_1, I_2) = \frac{\sum (i_1 - \bar{i}_1)(i_2 - \bar{i}_2) \mathcal{H}(i_1, i_2)}{\sqrt{(\sum (i_1 - \bar{i}_1)^2 \mathcal{H}(i_1, i_2)) \cdot (\sum (i_2 - \bar{i}_2)^2 \mathcal{H}(i_1, i_2))}} \quad (3.26)$$

The sums run over all intensity levels i_1 and i_2 while \bar{i}_1 and \bar{i}_2 are the average intensities in I_1 and I_2 , respectively.

If both images are sampled at the identical grid points, the histogram can be constructed very efficiently.

If an optimal transformation is only searched for in the translations space, the linear interpolation can be accelerated by an elegant trick based on the joint histograms.

Without loss of generality, let's assume that $I_2(\vec{q})$ is translated for a vector $\vec{p} = (p_1, p_2, p_3)$, where $p_1, p_2, p_3 \in [0,1)$. A tri-linearly interpolated translated image $I_2^{\vec{p}}(\vec{q})$ is calculated as follows:

$$\begin{aligned}
 I_2^{\vec{p}}(\vec{q}) &= I_2(\vec{q} + \vec{p}) \\
 &= \sum_{f_1=0,1} \sum_{f_2=0,1} \sum_{f_3=0,1} C_{\vec{p}}(\vec{f}) I_2(\vec{q} + \vec{f})
 \end{aligned}
 \tag{3.27}$$

Where the vector $\vec{f} = (f_1, f_2, f_3)$ in the sums is over the eight grid points nearest to \vec{p} (see Figure 44). Every pixel \vec{q} in the interpolated image $I_2^{\vec{p}}$ is constructed as an average of the eight pixels nearest to \vec{q} in the original image I_2 , weighted by

$$C_{\vec{p}}(\vec{f}) = \prod_{i=1}^3 p_i^{f_i} \cdot (1 - p_i)^{1-f_i}
 \tag{3.28}$$

Consequently, the joint histogram of an image I_1 and the interpolated image $I_2^{\vec{p}}$ can be calculated directly from eight histograms based only on non-interpolated images (\vec{f} are grid points by definition):

$$\mathcal{H}_{I_1, I_2^{\vec{p}}} = \sum_{f_1=0,1} \sum_{f_2=0,1} \sum_{f_3=0,1} C_{\vec{p}}(\vec{f}) \mathcal{H}_{I_1, I_2^{\vec{f}}}
 \tag{3.29}$$

Since a histogram is approximately two orders of magnitude less data than an image, the histogram interpolation eq. (3.29) is considerably faster than interpolation of an image, vastly accelerating the alignment procedure.

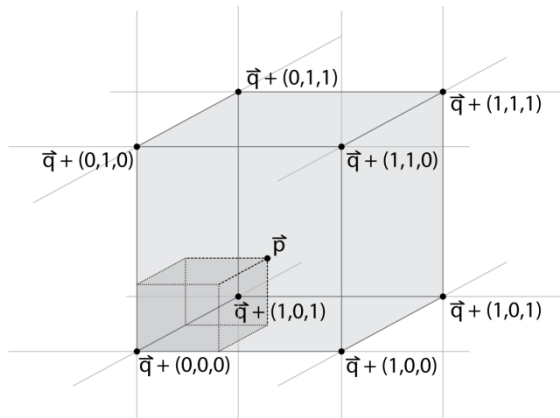


Figure 44: Arbitrary vector \vec{p} and its eight nearest grid points.

3.5.3.5 Early approaches to LSFM image alignment

In an ideal LSFM, the following assumptions about the multiple-view images can be made:

1. Each image is sampled precisely (beyond the limit defined on page 80) according to (3.11),
2. The vectors \vec{i} , \vec{j} and \vec{k} in (3.11) are mutually perpendicular and their lengths are precisely known,
3. The angular position of the rotary stage is precisely known,
4. The axis of rotation is parallel to one of the vectors \vec{i} , \vec{j} or \vec{k} .

Spatial relations between the different views are thus known, apart from their relative translation (\vec{r}_0). The preprocessing step (section 3.5.2) undoes the rotation, while the translation parameters are searched for by image alignment. If m is the number of the chosen reference image, then the transformation $\mathcal{M}_{m \rightarrow n}$, that maps the preprocessed image n onto a reference image m , has the following form:

$$\mathcal{M}_{m \rightarrow n}(\vec{q}) = \vec{q} + \vec{q}_n \quad (3.30)$$

The translation vectors in the early algorithms were optimized either by *phase correlation* [80,90,133] or by iterative similarity optimization.

Translation-only based image alignment is relatively quick (compared to alignment models requiring image interpolation). For example, the calculation of a single similarity metric of two 400 megabyte images (688x524x581 16 bit voxels) takes approx. 270ms on a four core 3 GHz PC¹⁶. The translation across a distance of 12.2 pixels (5, 11 and -2 along x, y and z, respectively) is optimized by a regular step linear search algorithm in 23 steps, directed by 57 metric evaluations. These metric evaluations already include translations corresponding to eight grid points nearest the metric maximum. The sub-pixel evaluation to arbitrary precision (allowed by linear interpolation) can, therefore, be performed without additional metric evaluations in less than 1 second using the histogram interpolation method described on page 84. The total time required for the alignment of two such images is, therefore, approximately 16 seconds (plus the time required to read the two images into RAM). Aligning a set of 8 views (400 MB each) therefore, takes approximately 2 minutes. There is some overhead for preprocessing, filtering, loading/saving and crude fusion, but the total procedure still stays in the range of minutes.

Unfortunately, this approach gave satisfactory results only sporadically. In most cases, a certain degree of misalignment was apparent (see Figure 45). While the alignment was acceptable in those parts of the images with high information content (bright areas, many objects), the divergence between different views became especially evident at the edges of the fused image. All this indicated that the four assumptions listed above did not hold. The preprocessed single views were obviously deformed by a more complex transformation than translation eq. (3.30).

¹⁶ Provided that both images fit into PC's RAM simultaneously.

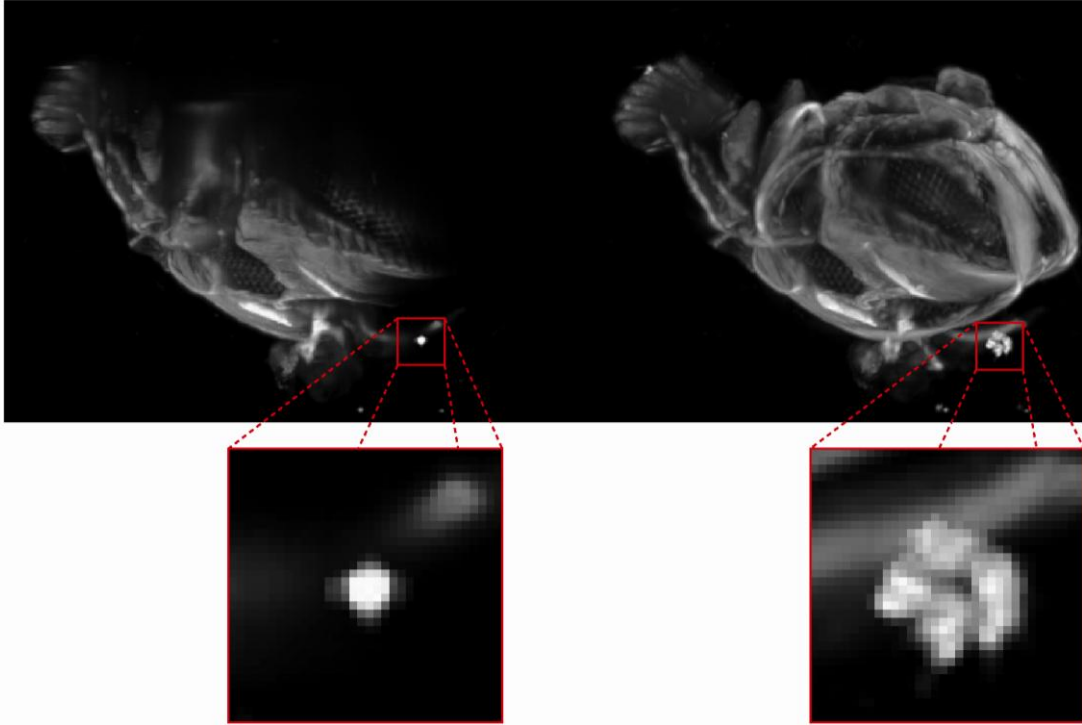


Figure 45: An example of image artifacts due to inaccurate image alignment. The left image shows a single view of an adult drosophila head (SPIM, Zeiss Achromat 10x/0.3W objective lens, autofluorescence signal, 488nm illumination, detection above 488nm) and the right image the fusion of 6 views aligned using the translation-only based model. While the center of the image shows an acceptable fit, volumes far-off from the center are visibly misaligned. The zoomed inset shows one such object positioned axially far from the head that gets effectively multiplied due to image misalignment. A more sophisticated model than translation is clearly required.

The problem often manifested itself in totally misaligned single views, i.e. a match between views after the alignment was not significant in any part of the image. The most likely cause for that is the fact, that the single views were so deformed (rotated, sheared and scaled) against each other that the correct maximum of the similarity metric completely leveled out and was thus hardly detectable. Equation (3.30) obviously does not adequately describe the relation between the different preprocessed single views.

3.5.3.6 Affine transformation image alignment

As demonstrated in previous section, more general coordinate mappings $\mathcal{M}_{m \rightarrow n}$ than translation (3.30) must be considered for alignment of LSFM images. While translation corresponds to a zero-order coordinate mapping, first order (i.e. linear) mapping is an affine transformation:

$$\mathcal{M}_{m \rightarrow n}(\vec{q}) = T_n \cdot \vec{q} + \vec{q}_n \quad (3.31)$$

where T_n is a 3x3 transformation matrix. Such a transformation can describe an arbitrary combination of elementary linear transformations: scaling, shear and rotation around an arbitrary axis. It is defined by 12 parameters (9 for T + 3 for \vec{q}_n).

LSFM image deformations

As sketched above, the goal is to align the single views with a precision that is better than approximately half of the optical resolution (section 3.5.3). However, as a rule of thumb, the

voxel pitch can serve as a rough guess for required precision¹⁷. Since a normal LSFM image consists of 1000-2000 voxels along each dimension, the precision of one voxel is achieved, if the residual (i.e. after the alignment) image deformation matrix T differs from an identity matrix by less than approximately 10^{-3} in each element.

In order to gain further insight into LSFM image deformations, all possible causes have to be identified. They are listed below in the order of assumed importance:

1. Tilted axis of rotation

An LSFM is normally constructed such that its axis of rotation is parallel to one of the image edges. In order to undo the mechanical rotation of the specimen and to bring all the views into the same orientation, the orientation of the axis of rotation has to be known with a precision that is better than $\sin^{-1}(10^{-3}) = 0.06^\circ$. This level of accuracy is hardly attainable on the hardware level, especially in the absence of a method that measures the orientation of the axis and fine-tunes it. The axis of rotation¹⁸ is therefore usually not sufficiently parallel to the image edge, which adds two free parameters to the image transformation model.

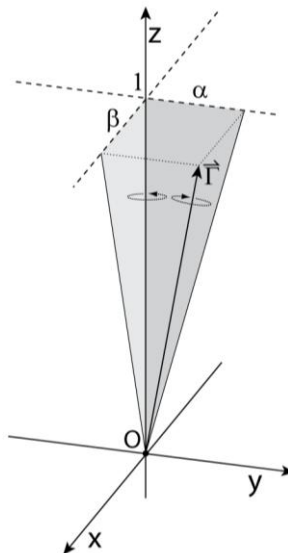


Figure 46: Tilted axis of rotation. Specimen rotation in the microscope is undone by counter rotation of the images. If the real axis of rotation $\vec{\Gamma}$ is tilted against the assumed axis of rotation \hat{z} (circular arrows), the counter-rotation around an incorrect axis will result in deformed images. Tilt in this diagram is vastly exaggerated; α and β are normally in the range of 10^{-3} .

Let us assume that the \hat{z} axis is parallel to one of the image edges and almost parallel to the axis of rotation $\vec{\Gamma}$. The axis of rotation almost parallel to \hat{z} can be constructed by rotating a unit vector parallel to \hat{z} for small angles α and β around \hat{x} and \hat{y} , respectively. In the case of $\alpha, \beta \ll 1$ this axis of rotation is approximately $\vec{\Gamma} \approx (\alpha, \beta, 1)$ (see Figure 46). The transformation operator from the image's coordinate system into a coordinate system, in which the axis of rotation is parallel to \hat{z} , is:

¹⁷ At least two voxels are acquired across the diameter of an Airy disk in a standard LSFM. Alignment precision of one voxel is therefore generally more strict than precision of half of the optical resolution.

¹⁸ It is important to realize that orientation of the image relative to the rotation axis depends also on the camera orientation. In some LSFM implementations, camera was freely rotatable around its optical axis. This was normally much more significant source of axis tilt than an eventual mechanical disorientation of the rotary motor.

$$R_\psi = \begin{bmatrix} 1 & 0 & -\alpha \\ 0 & 1 & -\beta \\ \alpha & \beta & 1 \end{bmatrix} \quad (3.32)$$

The rotation by an angle γ around the \hat{z} axis is described by the rotation matrix:

$$R_\gamma = \begin{bmatrix} \cos \gamma & -\sin \gamma & 0 \\ \sin \gamma & \cos \gamma & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (3.33)$$

Before the specimen is imaged, it is first rotated in the microscope by an angle γ around the axis $\vec{\Gamma}$. To undo the rotation, the image is digitally rotated by an angle $-\gamma$, assuming that the axis of rotation was parallel to \hat{z} . The deformation due to the mismatch between $\vec{\Gamma}$ and \hat{z} is, therefore, described by the following affine matrix:

$$D_\alpha = R_\gamma^{-1} \cdot (R_\psi^{-1} \cdot R_\gamma \cdot R_\psi) \approx \begin{bmatrix} 1 & 0 & -\alpha A - \beta B \\ 0 & 1 & -\alpha B + \beta A \\ \alpha A + \beta B & \alpha B - \beta A & 1 \end{bmatrix} \quad (3.34)$$

where $A = \sin \gamma$ and $B = (1 - \cos \gamma)$ and the expression in parentheses describes the rotation around the tilted axis $\vec{\Gamma}$. Since R_ψ is a rotation matrix, its inverse equals its transpose. The deformation D has a similar form as R_ψ . It, therefore, describes a small tilt in the orientation, which now depends also on the angle of rotation γ . The object in the preprocessed images will, therefore, seem to undergo precession as the rotation angle γ changes.

Using a method for characterization of LSFM image deformations described in section 3.5.3.7, α and β were usually measured to be in the range of 1-2°. Normally, the rotation around the optical axis is larger since it is additionally affected by the orientation of the camera.

2. Z-motor tilt

In all LSFM implementations up-to-date, three dimensional images are recorded by moving the specimen through the light sheet. A problem arises if this movement is not precisely (within the 0.06° limit, see above) parallel to the optical axis of the detection system (Figure 47).

The direction of the optical axis of the detection lens is normally designated as the \hat{z} axis. If the z-motor moves along a tilted axis $\vec{k} = (-\alpha, -\beta, 1)$ this causes a movement along the \hat{x} and \hat{y} directions whenever the specimen is translated along \hat{z} . This causes a shear in the images, which is described by the following two-parameter deformation matrix:

$$D_m = \begin{bmatrix} 1 & 0 & \alpha \\ 0 & 1 & \beta \\ 0 & 0 & 1 \end{bmatrix} \quad (3.35)$$

Normally, α (rotation around a horizontal axis) is less than 0.06° and can be neglected. On the other hand, β (rotation around a vertical axis) is on average around 1°.

3. Light-sheet tilt

Similarly, if the light sheet is not precisely perpendicular to the optical axis (\hat{z} axis), then tilted slices of a specimen are recorded (Figure 47c). Different parts of a single $\hat{x}\hat{y}$ slice are acquired in different \hat{z} positions, which again results in sheared images. The resulting deformation matrix has the following form

$$D_l = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \alpha & \beta & 1 \end{bmatrix} \quad (3.36)$$

If the light sheet normal is parallel to $\vec{h} = (-\alpha, -\beta, 1)$. By the method described in section 3.5.3.7 it has been determined, that α and β normally lie in the range of 0.5° .

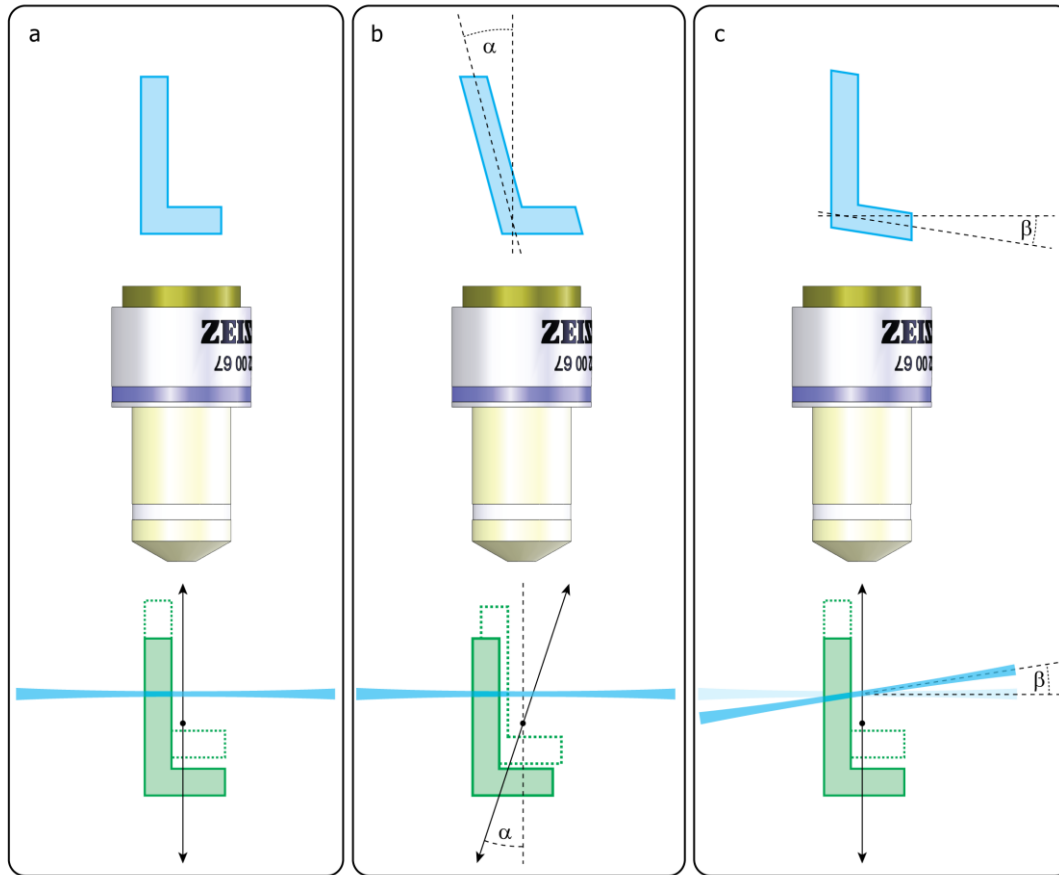


Figure 47: Sources of skew in LSFM images. **a** Optimal LSFM operation: the object is translated parallel to the optical axis and the light-sheet is precisely perpendicular to it. Image reflects the object well. **b** Object is translated along a tilted axis; image is laterally skewed. **c** Light-sheet is not precisely perpendicular to the optical axis; image is axially skewed. Real LSFM implementations exhibit combination of both kinds of skew. Tilts α and β in the schematics are exaggerated; they are normally in the range of 10^{-2} .

4. Axial/lateral scale mismatch

Lateral sampling ($\hat{x}\hat{y}$ plane) is defined by the magnification of the detection setup (section 2.2.1) and the image sensor pixel grid, while sampling along \hat{z} is determined by the translation of the specimen between consecutive slices. This can result in a mismatch¹⁹ between the sampling rates along lateral and axial directions, either due to inaccurate translation steps or inexact detection magnification. In any case, the deformation matrix has a form

¹⁹ In our experience, this mismatch is in range of <1% for most lenses, while it can get as high as 4% if the lens are not used as they were designed to, e.g. air lens used to observe a water immersed specimen through a glass window.

$$D_s = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 + d \end{bmatrix} \quad (3.37)$$

where d is a correction factor. It is usually in range of 0.01, but it can be considerably larger (up to 0.05) if the detection objective lens is used in a configuration it was not designed for.

5. Wide-field aberrations

Geometric aberrations in the detection optics are expected to contribute only a marginal degree to image deformation, however they cannot be completely ruled out. The only linear geometrical deformations, which retain a circular symmetry of the detection lens are scaling, which is included above (4), and rotation (1). All remaining deformations, e.g. radially varying magnification (i.e. barrel/pincushion distortion), are nonlinear in nature and cannot be described by an affine transformation (3.31). Fortunately, they have not presented a problem in LSFM image fusion until now.

Searching the space of all 12 parameters that define an affine transformation (3.31) is computationally so demanding, that it can hardly be used for the alignment of LSFM views. For example, a four core 3 GHz PC requires approximately 52 seconds to calculate a single cross-correlation of two 370 megabyte images. This is approximately 200 times more than the time needed for a translation-only based normalized cross-correlation as described on page 84. The setback is mostly due to the fact, that the image values that are not defined by the original grid of a transformed image, must be calculated. A linear interpolation was used in the example above. Using a simpler interpolation, e.g. a nearest neighbor, can improve the speed at a significant loss of accuracy and robustness.

Another disadvantage of a full affine alignment is the high dimensionality of the parameter space. This means that the optimization will converge significantly slower than if a more limited number of parameters is searched for. As a rule of thumb, doubling the number of free parameters squares the number of metric evaluations required to find the minimum. The reality is usually not so bleak, however at least several hundred evaluations are normally required to align two moderately displaced (around 20 voxels) LSFM views. The total time required to align two 400Mb views is therefore around 4-5 hours, and 30-40 hours for a set of 8 views.

A pyramid approach was also tried to improve the convergence speed. The central idea of pyramid alignment is to obtain rough estimates of optimal parameters by aligning downsized images and then improving the guess by repeating the procedure with less reduced images (Figure 48). In most real LSFM registration applications, pyramid approach provided 2 - 2.5× faster convergence. Unfortunately, 10-30 hours are still required to register standard multiple-view set using affine transformation.

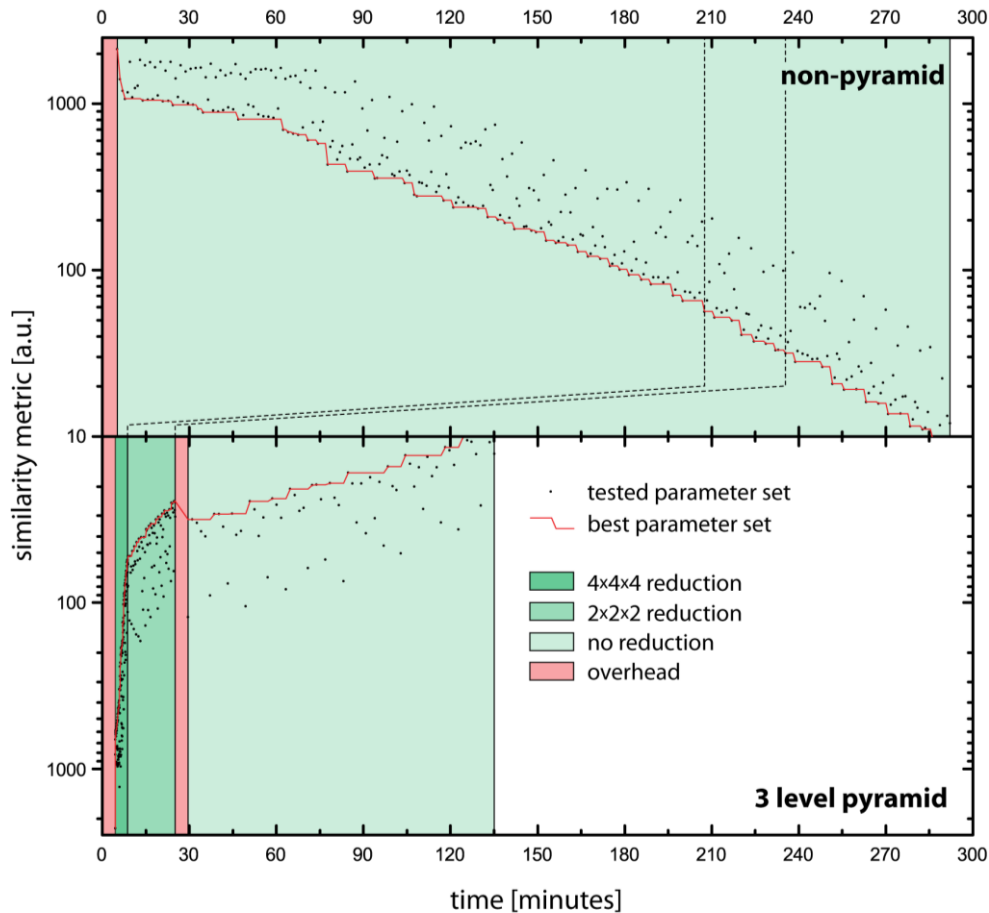


Figure 48: Speed improvement by pyramid alignment of LSFM views. Two 400 megabyte images (approx. $770 \times 380 \times 680$ voxels) were aligned using a non-pyramid affine alignment (top graph) and 3 level pyramid (bottom graph). Computation time scales linearly with the image sizes; 1.41s, 11.4s and 91.8s per iteration were required in the first (64× reduced), second (8× reduced) and third (not reduced) level of the pyramid, respectively. Dashed lines connect points in the both graphs with the same value of the similarity metric. First level of the pyramid reached in 7 minutes the same score as non-pyramid alignment in 209 minutes – almost 30× quicker. However, differences get smaller at lower levels. Second level is only 2× as efficient as the non-pyramid algorithm while final refinement on the non-reduced level takes even longer than corresponding refinement by the non-pyramid algorithm. Total improvement is in range of 2×.

As discussed above, LSFM image deformations can be described by 7 parameters while a general affine transformation matrix includes 9 independent parameters. The search space could therefore be reduced by two dimensions, which is estimated (in best case) to half the number of required metric evaluations and the necessary time. Such an optimization would still take an unreasonably long time and the improvement did not seem interesting enough for us to implement it. The search space could be limited further by constraining the optimization to the parameters that are expected to have the greatest effect on the LSFM image deformation. A five-parameter alignment of LSFM data, constrained to translation (3 parameters) and orientation of rotation axis (2 parameters; see above), has been shown to give reasonable results [157].

3.5.3.7 Characterization of LSFM image deformations

An alternative to a full affine image registration as described in the previous section is to measure the deformation matrices T_n of equation (3.31) prior to the experiment. Assuming that LSFM imaging deformations are determined only by the microscope, the specimens can be exchanged

without modifying the T_n . Subsequent images are then corrected by the a priori measured deformation matrix and fused by the rapid translation-only algorithm described in section Early approaches to LSFM image alignment.

Deformations can be measured using an arbitrary specimen that produces robustly registrable images. The fluorescent beads homogeneously dispersed in a block of transparent gel have proven to be a very useful specimen for the the diagnostics of LSFM microscopes. They provide a high signal-to-noise ratio, low photo bleaching and images rich in the high-frequencies needed for an accurate registration. Such a specimen is almost transparent, thus the views taken along two arbitrary directions always exhibit sufficient common information for an exact registration. Pair-wise registration (page 80) is therefore not required; all views can be registered with a single, arbitrarily chosen reference view. Last but not least, such a diagnostic specimen is easily prepared and can be stored over long periods of time.

On the other hand, an image of dispersed fluorescent beads lacks low-frequency information. When two such images are registered, the similarity metric is a flat function with a large number of local maxima (whenever any bead in the first image overlaps with any bead in the second image) that can distract the optimization away from the global maximum. The registration will, therefore, work well only if a reasonably good starting point is provided. Furthermore, if affine based image alignment (previous chapter) is used to obtain deformation matrices, the analysis can take several days to complete.

Alternatively, each of the images of fluorescent beads can be described by a list of coordinates of the beads in the image. An image with a size in range of a gigabyte is thus reduced to a kilobyte sized numeric list. The optimization of the transformation parameters is then done directly on this coordinate list. This process is several orders of magnitude faster than performing it on the whole images.

Fluorescent beads produce images with a high contrast and they can, therefore, be segmented simply by a threshold, i.e. detecting areas in the image with intensities above a certain intensity level. Connected bright areas are then labeled (e.g. `bwlabel` function in MatLab, see also [158]) and the coordinates of the beads in the images are calculated as centroids of labeled regions. Each three-dimensional image is thus turned into a list of coordinates representing positions of the beads in the set of images. The number of the beads in a set of images (normally several hundred) is much higher than the number of the deformation parameters. Any unsystematic errors produced by the simple segmentation process described above will therefore be averaged out in the processing that follows and will not affect the final result. Furthermore, due to the high number of beads, the registration precision is expected to be in the sub-pixel range.

The segmentation reduces the amount of data by at least five orders of magnitude (i.e. 100,000). The result is a list of bead coordinates for the superset of N images recorded along N different directions.

$$\begin{aligned}\vec{v}_i^1 &= (x_i^1, y_i^1, z_i^1) \wedge i \in \{1, \dots, n_1\} \\ \vec{v}_i^2 &= (x_i^2, y_i^2, z_i^2) \wedge i \in \{1, \dots, n_2\} \\ &\vdots \\ \vec{v}_i^N &= (x_i^N, y_i^N, z_i^N) \wedge i \in \{1, \dots, n_N\}\end{aligned}$$

Here \overrightarrow{v}_i^1 and \overrightarrow{v}_i^2 refer to lists of coordinates obtained from the first and the second view of the same bead specimen. The values n_i are the numbers of detected beads in the respective lists.

Although they all represent the same specimen, the lists obtained from any two views can vary significantly. The number of detected beads is normally different, beads are listed in an arbitrary order and not all beads are visible in each view. It is, therefore, not clear which entry in each of the lists corresponds to the same bead in the specimen. Coordinate sets still have to be registered. Again, the affine transformation model is used:

$$\overrightarrow{u}_i^k = T_k \cdot \overrightarrow{v}_i^k + \overrightarrow{t}^k \quad (3.38)$$

where $k \in \{1, \dots, N\}$ is the view number. Usually one of the lists (e.g. the first view) is chosen as the reference (r) to which other views are registered.

To register the images, a merit function $\chi(\hat{A}, \vec{t})$ mimicking the image cross correlation (page 82) is constructed. The merit function first calculates the distances between all coordinates in the transformed list (\overrightarrow{u}_i^k) and all coordinates in the un-transformed reference list (\overrightarrow{v}_j^r). So if list k contains the coordinates of n beads and reference list r the coordinates of m beads the result is a list of $n \times m$ distances. These distances d_{ij}^2 are then weighted by a Gaussian function (good approximation of a linear PSF profile) that penalizes the distances that are considerably larger than e.g. the lateral extent of the point spread function:

$$f(d_{ij}^2) = e^{-\frac{d_{ij}^2}{2\sigma^2}} \quad (3.39)$$

Standard deviation σ describes the cut-off value between short distances that will receive a weight close to 1 and the long distances, receiving weight closer to 0. The exact nature of f is not crucial as long as it converges to 0 faster than linearly. The sum of all $f(d_{nm}^2)$ is the result of the merit function χ_k :

$$\chi_k(T_k, \overrightarrow{t}^k) = \sum_{i=1}^n \sum_{j=1}^m f\left(\left((T_k \cdot \overrightarrow{v}_i^k + \overrightarrow{t}^k) - \overrightarrow{v}_j^r\right)^2\right) \quad (3.40)$$

The function χ_k estimates how many coordinates in the list \overrightarrow{v}_i^k have an imminent vector in the list \overrightarrow{v}_j^r after the later is transformed as defined by T_k and \overrightarrow{t}^k . This merit function is mainly a flat hyper-surface with a number of local maxima that can distract the optimization process. A starting point within the catch-radius of the global maximum is, therefore, required (however, the catch radius can be initially made larger by increasing σ in (3.40)).

Different views are rotated relative to each other around. A good starting approximation of T_k is, therefore, a rotation matrix corresponding to the orientation of the view k relative to the reference view r . Most of the coordinates in the list k have a counterpart in the list r that corresponds to an image of the same fluorescent bead. If the vectors \overrightarrow{v}_i^k and \overrightarrow{v}_j^r form a pair then an approximate translation vector can be calculated from the relation $\overrightarrow{v}_j^r = T_k \cdot \overrightarrow{v}_i^k + \overrightarrow{t}^k$ with an approximate T_k .

One way of finding pairs \overline{v}_i^k and \overline{v}_j^r that correspond to the same bead is to randomly select a coordinate from each of the two lists, calculate approximate T_k and \overline{t}^k as sketched above and feed them into the merit function (3.40). Matching pairs can be identified by an extraordinary high merit. If this procedure is repeated sufficiently often, the probability of identifying at least one matching pair will approach unity²⁰. A more systematic and robust approach is to check a small number of coordinates from the list k against all coordinates in the list r . With two lists of 100 coordinates each, one coordinate from the first list is tested against all coordinates in the second list in approx. 360 milliseconds on a 3GHz PC.

Once a starting point near the global minimum of the merit function $\chi_k(T_k, \overline{t}^k)$ is determined, elements of T_k and \overline{t}^k are optimized by a continuous function optimization algorithm (e.g. *fminsearch* in *MatLab* or *Fit* in *Mathematica*).

The procedure determines the deformation part of the linear mappings (defined by T_k) that optimally relate the views to each other. If the mechanical parts of the LSFM are so stable that the deformation parameters do not change when the specimen is rotated or exchanged, this information can be used to correct all subsequently recorded images. The preprocessing (section 3.5.2) part is replaced by an affine transformations based on the acquired T_k to transform the views into a common orientation. This allows a rapid translation-only alignment of LSFM views (Chapter 3.5.3.5).

EMBL's LSFM implementations do seem to demonstrate a sufficient stability²¹. The stability is assessed by imaging the diagnostic bead specimen multiple times and by comparing the matrices produced every time. An example of such experiment is shown in Figure 49.

Alternatively, the marker based alignment can be used also by mixing the fluorescent beads into the block of gel containing the specimen. The beads are then segmented and their positions registered. The resulting parameters are then used to fully align the images. Unfortunately, the ensuing microscopic images will demonstrate a host of bright spots produced by the beads. These spots are usually brighter than the actual fluorescent specimen that one intends to investigate, they are difficult to remove from the images and present a serious nuisance. Two ways were proposed to avoid that problem: i) the beads used for alignment are fluorescent at a different excitation/emission wavelengths or ii) the beads are spatially separated from the specimen, e.g. by including them in a layer of the agarose block below or above the specimen.

²⁰ It was recently revealed to me by Dr. Pavel Tomancak (Max Planck Institute of Molecular Cell Biology and Genetics, Dresden), that this Monte Carlo approach is commonly used in model fitting applications and is known as RANdoM SAmple Consensus (RAMSAC). The mathematical framework at the core of the RAMSAC theory also includes methods to optimize the number of required random samples and the way in which they are picked [174].

²¹ There were personal reports from some other laboratories that their LSFM's rotary stage was not sufficiently stable and that a five parameter (translation + axis of rotation) image alignment was indispensable.

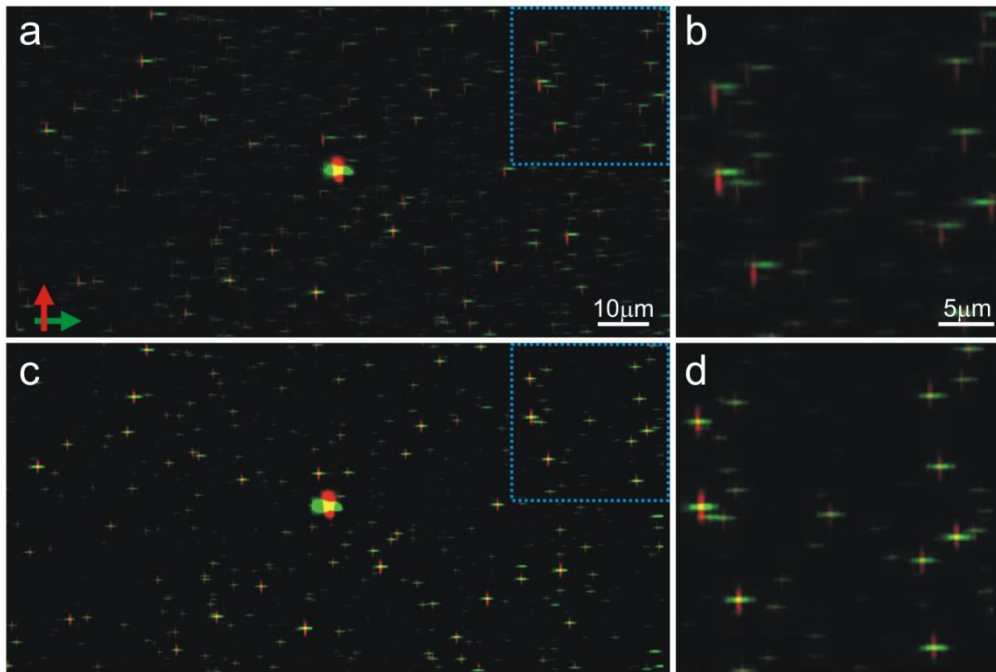


Figure 49: Effect of image deformation correction on image alignment. The two multi-view SPIM images of fluorescent beads were acquired along the two different directions shown in red and green. The red and green arrows in **a** indicate the orientation of the detection axes in the respective views of the same colour. **a** The result of a translation image alignment without a prior deformation correction. **b** The magnification of the region in the upper right of **a** indicated by the blue square. **c** and **d** The result of a translation alignment after the image stack was corrected by an affine transformation. The real space images are maximum intensity projections. The original data set has a size of 1024x1344x512 voxels. The image stacks were recorded with a Zeiss Achroplan 40x/0.8W, the excitation wavelength was 0.488 μm , the emission was recorded above 0.488 μm . The scale bars in **a** and **b** are also valid for **c** and **d** respectively.

3.5.4 Final image fusion

The final fusion refers to the concluding part of the full image fusion algorithm. In this step, the information contained in the input images is finally combined into a single image. Such an image must maintain the “good” data while the “bad” data is replaced with complementary data from alternative views.

As mentioned in section 3.2, the multiple-view image fusion algorithms aim at two fundamentally different goals: a) completing the image of opaque specimens and b) improving the image resolution, which will also make it more isotropic. Unfortunately, no robust final fusion algorithm that reliably produces good results regardless of the input images’ properties was reported so far. While the previous steps in the image processing pipeline work well with differently transparent specimens, a final fusion algorithm that suits the properties of the input images must be selected. Extended and opaque specimens require final fusion algorithms that are able to calculate and to take the local image quality of each of the views into account.

A number of methods that aim at either of the two goals were proposed in the past. Some of the most commonly used will be evaluated in the following two sections, together with a novel algorithm developed at EMBL.

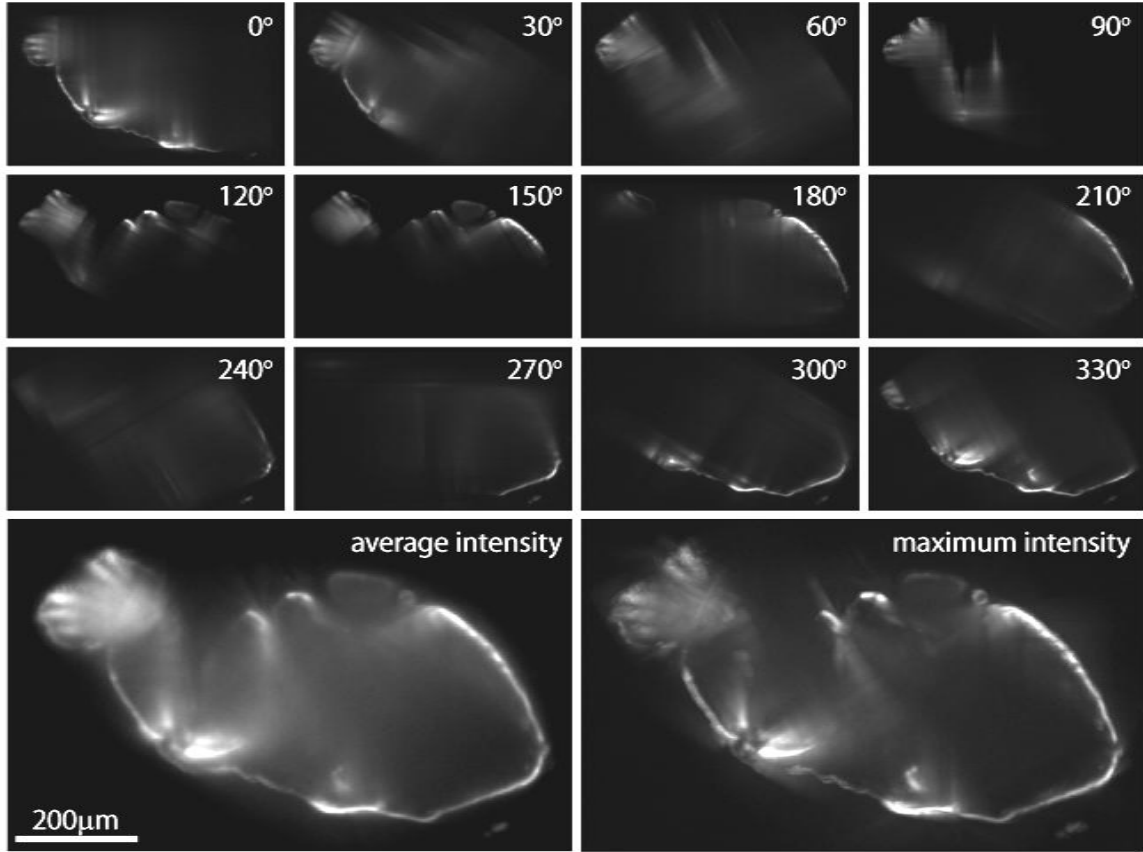


Figure 50: Average and maximum intensity fusion methods in case of non-transparent specimen. Twelve views of an adult *Drosophila melanogaster* fly head (top three rows) were fused using average intensity (bottom left) and maximum intensity (bottom right) fusion methods. Images show cross-section through the middle of the head. While each of the input views reveals only a very limited part of head's surface, both fusions are complete. Average intensity suffers from intensity decrease due to the fact, that every part of the surface is only visible in a limited number of views. Image's brightness was increased for approx. 3× to match the intensity of the other images, which creates the impression of amplified blurred background. Maximum intensity, on the other hand, preserves anisotropic artifacts from the individual input views (e.g. bright streaks radiating from intense objects). Fruit fly's autofluorescence was imaged with SPIM, Fluar 5x/0.25 objective lens, illuminated with wavelength 0.488µm through cylindrical lens only, emission above 0.488µm was detected. Maximum parallel projection of the stack is shown in Figure 69.

3.5.4.1 Mosaicing algorithms

Mosaicing (or stitching) algorithms concentrate solely on creating a more complete image from a set of images. Each image covers a part of the specimen but overlaps with at least one other image. All algorithms in this category create fused image by means of an arithmetic average of the input views [159]:

$$M(x, y, z) = \frac{\sum_i w_i(x, y, z) I_i(x, y, z)}{\sum_i w_i(x, y, z)} \quad (3.41)$$

where the sums extend over all N input views $i = 1 \dots N$ and $w_i(x, y, z)$ are the local weighting factors, which define how much each of the views i contributes to the final fusion at each point (x, y, z) .

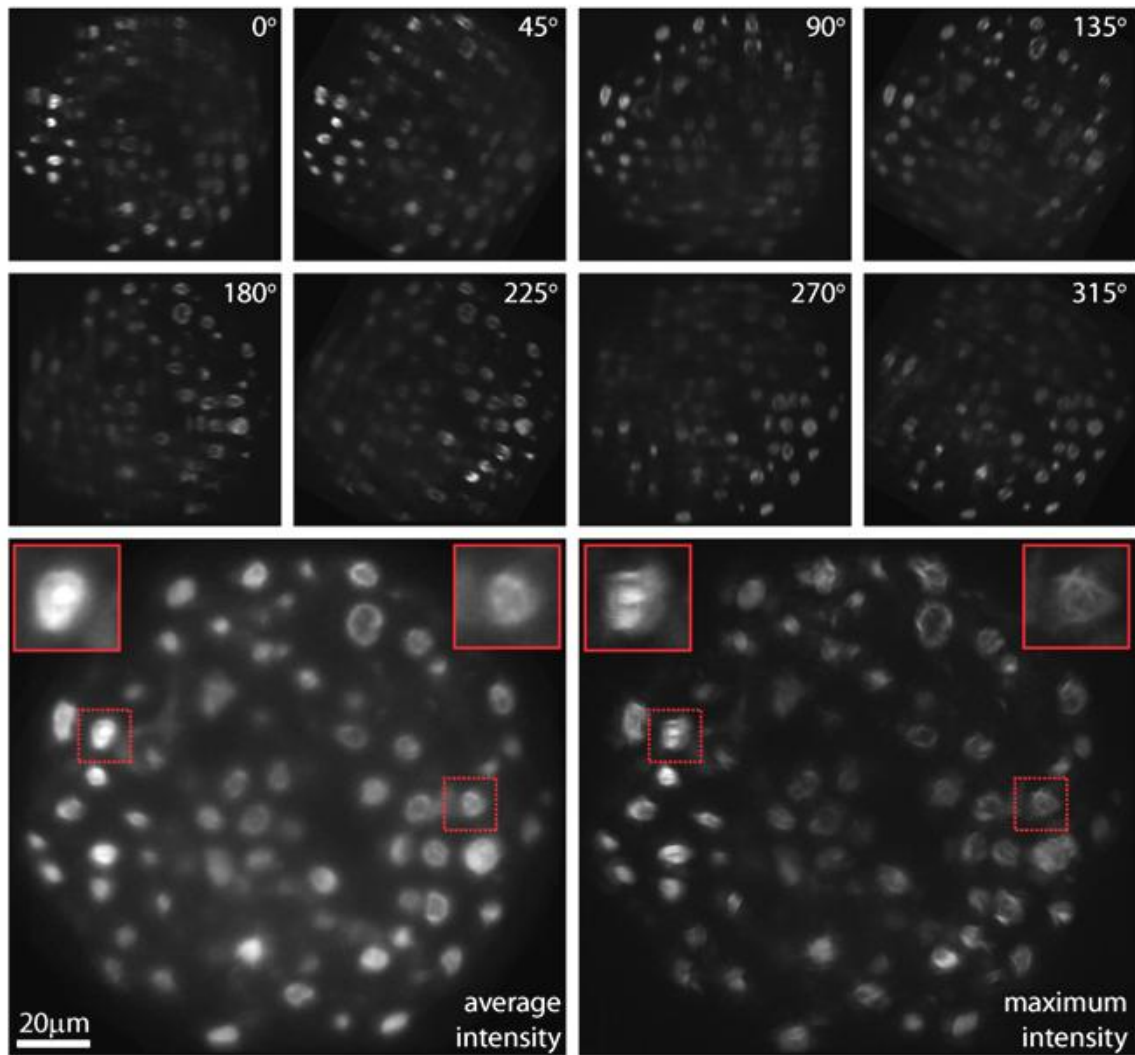


Figure 51: Average and maximum intensity fusion methods in case of semi-transparent specimen. Eight SPIM views (top two rows) of a large cellular spheroid of BXP3 human pancreatic cancer cells labeled with DRAQ5 are fused using average intensity (bottom left) and maximum intensity (bottom right) methods. Images show cross-section through the middle of the cell-body. Again, average intensity fusion attenuates bright objects for approx. $3\times$ since every object is well visible only on a small number of the views and dark otherwise. This creates impression of amplified background. Maximum intensity fusion keeps high intensity of the input views, but it also preserves the anisotropic artifacts from individual views, e.g. orientation of individual PSFs from contributing views is clearly visible in the fusion. Both methods retain high level of background from the blurred parts of the views. The insets show magnified two selected nuclei. The original data set has a size of $1024 \times 1344 \times 256$ picture elements. The image stacks were recorded with a Zeiss Achroplan 40x/0.8W, the excitation wavelength was $0.488\mu\text{m}$ and the emission was recorded above $0.520\mu\text{m}$. The sample was provided by Marco Marcello (DKFZ-Heidelberg) and Francesco Pampaloni (EMBL).

The simplest method is a non-weighted average: $w_i(x, y, z) = 1$. It produces an image that covers the whole specimen more homogeneously than any single view alone (Figure 50, bottom left). Non-weighted average treats all views equally and, therefore, does not discriminate the “good” parts of views against the “bad” parts. This manifests itself in two ways: i) smear, present in any of the views, will be unselectively incorporated in the fused image and ii) parts of the image that are only bright in some of the views will look muted in the fusion (Figure 50 and Figure 51). Better results are achieved if $w_i(x, y, z)$ reflects the quality of the view i in the vicinity of the point (x, y, z) .

3.5.4.2 Intensity weighted fusion

The image degradation normally manifests itself in two ways: intensity attenuation and blur. Due to the former, the local quality in each of the views can be estimated using its local intensity relative to the other views: $w_i(x, y, z) = I_i^p(x, y, z)$. The exponent p determines the influence of the intensity on the weight; $p = 0$ produces non-weighted average images and $p \gg 1$ will essentially produce a fusion by picking the brightest pixels: $M(x, y, z) = \max_i(I_i(x, y, z))$. Both extremes are shown in Figure 50 and Figure 51.

The maximum-intensity fusion produces images with higher intensity levels than a non-weighted average. However, it performs only marginally better in terms of blur-rejection. This can be explained by the fact, that the blur increases the intensity of dark regions around a bright object. Some areas of blurred images are therefore brighter than they are in their less blurred counterparts, and blur is thus incorporated into the fused image (Figure 51). Furthermore, anisotropic blur has a different orientation in each of the multiple views. Maximum intensity fusion combines blur from all views, producing star-shaped artifacts in the final image (Figure 51, inset). In non-transparent specimens, intensity attenuation (or even total obscurement) is the prevailing mechanism of image degradation (Figure 50). Multiple-view fusion images therefore suffer less from the poor blur-rejection properties of the fusion methods discussed above (Figure 51). The intensity-weighted average and maximum intensity fusion have both been proven useful for fusion of images of poorly-transparent specimens.

3.5.4.3 High-frequency weighted fusion

As mentioned above, image degradation manifests itself in two ways: intensity attenuation and blur. While the intensity-attenuation results in a homogeneous decrease in amplitude across the entire spectrum of the image, the blurring is equivalent to a decrease of the high frequency part of the spectrum, i.e. it is a low pass filter. The “good” parts of the image can, therefore, be more robustly identified by the high-frequency image content. A good choice for $w_i(x, y, z)$ is, therefore, a high-pass filtered image $I_i(x, y, z)$.

The original high-frequency weighted image fusion implementation based on formula (3.41) employed first and second order derivatives of the image intensity as a measure of the high-frequency content in the image [159]:

$$\begin{aligned} w_i^{d_1}(x, y, z) &= \|\nabla I_i(x, y, z)\| = \left\| \left(\frac{\partial I_i}{\partial x}, \frac{\partial I_i}{\partial y}, \frac{\partial I_i}{\partial z} \right) \right\| \\ w_i^{d_2}(x, y, z) &= |\nabla^2 I_i(x, y, z)| = \left| \frac{\partial^2 I_i}{\partial x^2} + \frac{\partial^2 I_i}{\partial y^2} + \frac{\partial^2 I_i}{\partial z^2} \right| \end{aligned} \quad (3.42)$$

In regularly spaced discrete coordinates of the images’ grids, the first and second intensity differential are calculated by convolving the images along every dimension with kernels $D_1 = [-1, 1]$ and $D_2 = [1, -2, 1]$, respectively [160]. The first intensity derivative in (3.42) is a vector and its norm is thus used in (3.42). Furthermore, the second derivatives along different dimensions can have opposite signs and can, therefore, cancel their contributions, even if their amplitudes are high. It is better to use modified differentiation formulas that add absolute amplitudes of the derivatives along each of the three dimensions:

$$\begin{aligned}
w_i^{D_1}(x, y, z) &= \left| \frac{\partial I_i}{\partial x} \right| + \left| \frac{\partial I_i}{\partial y} \right| + \left| \frac{\partial I_i}{\partial z} \right| = |I_i * D_{1,x}| + |I_i * D_{1,y}| + |I_i * D_{1,z}| \\
w_i^{D_2}(x, y, z) &= \left| \frac{\partial^2 I_i}{\partial x^2} \right| + \left| \frac{\partial^2 I_i}{\partial y^2} \right| + \left| \frac{\partial^2 I_i}{\partial z^2} \right| = |I_i * D_{2,x}| + |I_i * D_{2,y}| + |I_i * D_{2,z}|
\end{aligned} \tag{3.43}$$

Here, $*$ denotes convolution, $D_{1,x}$, $D_{1,y}$ and $D_{1,z}$ are the first order derivative kernels along the x, y and z axes, and $D_{2,x}$, $D_{2,y}$ and $D_{2,z}$ are the second order derivative kernels along the x, y and z axes, respectively. My numerical evaluations have shown that both weighting functions (3.43) identify high-frequency rich parts of the input views well (Figure 52, first row). However, they visibly amplify high-frequency noise in dim parts of the image. This can be explained by the fact, that derivative based measures (3.43), as defined by convolution kernels $[-1, 1]$ and $[1, -2, 1]$, are tuned to detect the amplitudes near the highest frequency present in the image, i.e. at the Nyquist frequency. On the other hand, the LSFM's PSF is normally covered by 3-5 voxels along each dimension (see section 2.2.3). This means that the "useful" image content is limited to the part of the image's spectrum corresponding to distances larger than 3-5 voxels, while functions (3.43) is sensitive to changes on a scale of 1-2 voxels. Such high frequencies are weak in LSFM images and are easily drowned in noise. This is most fatally demonstrated in the dim parts of an image (Figure 52, insets in first row).

A good workaround is to blur the weighting function $w_i(x, y, z)$. As both, differentiation and blurring, are described by convolution, they can be joined by convolving their respective kernels:

$$\begin{aligned}
D_1^m &= [-1 \quad 1] * \underbrace{[1 \quad \dots \quad 1]}_m = \underbrace{[-1 \quad 0 \quad \dots \quad 0 \quad 1]}_{(m+1)} \\
D_2^m &= [1 \quad -2 \quad 1] * \underbrace{[1 \quad \dots \quad 1]}_m = \underbrace{[1 \quad -1 \quad 0 \quad \dots \quad 0 \quad -1 \quad 1]}_{(m+2)}
\end{aligned} \tag{3.44}$$

where D_1^m and D_2^m stand for first and second order derivative kernels, respectively, blurred with a box kernel of length m . Again, the differentiation kernels (3.44) must be applied along every image dimension separately as in (3.43). It is evident from (3.44) that such elongated derivation kernels measure variations on larger scales and are better suited to detect the relevant high-frequency content in LSFM images. Numerical simulations have shown that kernels with sizes of m in the range of 3-5 voxels yield good results. This is no surprise considering that LSFM's PSF normal stretches across 3-5 voxels. In the case of a larger or a smaller PSF, the kernels (3.44) must be scaled correspondingly. The result of the blurring window size on noise amplification is demonstrated in Figure 52.

In the recent years, several noteworthy alternatives to the high-frequency weighted fusion methods described above were proposed. The local image entropy was suggested as a means for estimating the local image quality for the fusion of multiple exposure photographs [161]. An entropy weighted average fusion was also tested with LSFM images [129] and was shown to successfully discriminate "good" parts of the views against the "bad" ones.

The modeling of photon scattering in tissues [162] was also put forward as a way to determine local image quality in each of the LSFM views [143]. The images are first used to determine the outer shape of the tissue (i.e. mouse embryo) in each of the views. This information is then used to calculate the distance that the light had to transverse through the tissue to get from a given

point in each of the views to the detection objective lens. The image quality in that point is then estimated using a Monte Carlo simulation of light scattering. The simulated image quality is finally used for the weighting factors $w_i(x, y, z)$ of the formula (3.41), favoring the parts of the views that were closer to the detection objective. The method was only tested on a two-view system.

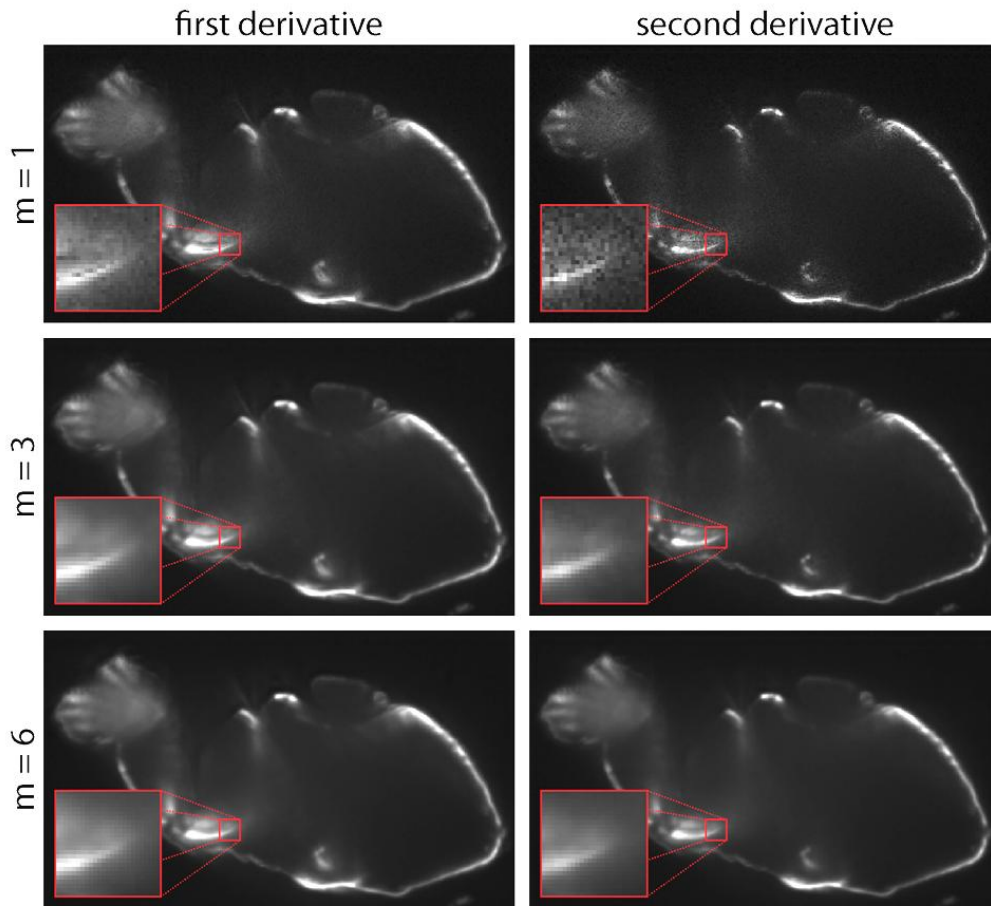


Figure 52: Derivative weighted fusion method. Eight SPIM views (see Figure 50) of a of an adult *Drosophila melanogaster* fly head were fused using a derivative weighted average. First order derivative (left column) and second order (right column) were tried as estimates of local image quality. Estimates were additionally blurred with a 3x3 (second row) and 6x6 (third row) box kernel to suppress the high frequency noise (see magnifications). Both derivatives seem to provide a good measure of local image quality. Good parts of individual views are incorporated into the fusions without the intensity loss or anisotropic artifacts suffered by non-weighted average and maximum intensity fusions (shown in Figure 50). For details about imaging conditions, see Figure 50.

All methods sketched above successfully fuse multiple views of opaque specimens into an image that is more complete than any single view. Furthermore, the methods are easily implemented and parallelized and computationally not very demanding. As a rule of thumb, the fusion of eight views never takes more than approx. 20 minutes; the final goal of simultaneous imaging and image fusion is almost within reach.

Though there were indications, that the weighted average fusion (3.41) could improve the resolution in the fused image with the right choice of weighting factors $w_i(x, y, z)$, this was never correctly demonstrated. Experiments and intuition show that these methods create an image fusion that resembles a fluorophore distribution that is better than any single view, however this

was never stringently. The following section discusses a group of more sophisticated and mathematically sound fusion methods that do improve image resolution.

3.5.4.4 Resolution improving methods

The following group of algorithms disregards the fact that image quality normally varies across the image and that different views image well different parts of the specimen. They deal predominantly with the challenge of improving the fused image's resolution and making it more isotropic, assuming a space-invariant image quality.

The linear image formation process is most generally described by equation (3.13). The images are normally expressed in a regularly sampled form $I_i(x, y, z)$ where the coordinates span a rectangular grid of m_i^x , m_i^y and m_i^z points along the x , y and z axes, totaling $M_i = m_i^x \cdot m_i^y \cdot m_i^z$ voxels (see page 78). The fluorescent object (more precisely three-dimensional fluorophore distribution) $O(x, y, z)$ that the images I_i represent is restored on a discrete grid of $L = l^x \cdot l^y \cdot l^z$ voxels. The discrete version of equation (3.13) takes the following matrix form:

$$\mathbf{g}_i = \mathbf{H}_i \cdot \mathbf{f} + \mathbf{n}_i, \quad i = 1, \dots, N \quad (3.45)$$

where N is the number of views, \mathbf{g}_i is a vector of length M_i consisting of stacked voxels of image I_i , \mathbf{f} is the object vector of length L , \mathbf{H}_i is the $M_i \times L$ blurring matrix of the view i and \mathbf{n}_i represents additive noise. The blur matrices \mathbf{H}_i the normally well-known (either calculated or measured PSF of the microscope). Set of equations (3.45) can be combined in a single linear equation:

$$\mathbf{g} = \mathbf{H} \cdot \mathbf{f} + \mathbf{n} \quad (3.46)$$

where \mathbf{g} is a vector of size $\sum_i M_i$ consisting of the voxels of all N images $\mathbf{g} = [\mathbf{g}_1, \mathbf{g}_2 \dots \mathbf{g}_N]$, \mathbf{H} is a $(\sum_i M_i) \times L$ combined blur matrix and \mathbf{n} is a noise vector of size $\sum_i M_i$.

The problem of obtaining an estimate of the fluorophore distribution \mathbf{f} that was likely to generate the images \mathbf{g} should be a standard linear inverse problem. Disregarding the noise, equation (3.46) can be inverted and has a unique solution if and only if $(\sum_i M_i) = L$. In the much more likely case of an overdetermined system $(\sum_i M_i) > L$, an approximate solution is obtained, for example, by minimizing the L_2 norm of the difference vector $\min_f \|\mathbf{H} \cdot \mathbf{f} - \mathbf{g}\|$. Unfortunately, such direct approaches yield poor results and generate non-continuous solutions that usually also violate the non-negativity constraint of fluorescence images $\mathbf{f} \geq 0$ [163]. This is easily comprehensible in the light of the fact, that due to the blurring with h , the $I_i(x, y, z)$ are low-pass filtered images of the original object $O(x, y, z)$. Frequencies above a certain limit are inevitably weaker than the high-frequency noise. High frequencies in the direct solution for \mathbf{f} are therefore governed predominantly by the noise \mathbf{n} and bear poor or no resemblance to the underlying fluorophore distribution. Power spectrum of the solution \mathbf{f} should therefore be limited to the low-frequency range.

It has to be understood that an exact solution of (3.46) is not required (and probably does not even exist) since \mathbf{g} is only measured with a certain (finite) precision. Any vector $\tilde{\mathbf{f}}$, which is expected to generate measurements $\tilde{\mathbf{g}}$ close enough to the measured \mathbf{g} such that the differences

can be attributed to the measurement noise \mathbf{n} , is a valid solution. The problem of finding a vector $\tilde{\mathbf{f}}$ that is likely to have generated measurements \mathbf{g} is therefore a matter of statistical estimation rather than a simple exercise in solving linear equations.

A number of image fusion algorithms were proposed for generating an approximate solution \mathbf{f} from a set of images \mathbf{g} (or equivalently from \mathbf{g}_i). Most algorithms do so by intuitively constructed mechanisms, but only one of them (discussed in section 3.5.4.13) is based on a sound mathematical framework. Some of the most noticeable such algorithms are explained in the following sections.

3.5.4.5 Frequency domain based algorithms

As mentioned before, microscope acts as a amplitude filter. It constrains the image spectrum to a narrower band along its axial direction than along its lateral directions. Multiple-view imaging complements reduced axial spectrum of one view with the lateral spectrum of complementary views (see also Figure 37). Therefore, many resolution-improving fusion algorithms are based on Fourier domain algorithms, which literarily fill the frequency space with the information from the complementary single views. These methods are conceptually and formally equivalent to the mosaicing methods described in section 3.5.4.1. However, the ‘‘mosaic’’ of individual views is assembled in the frequency space.

3.5.4.6 Amplitude-weighted frequency-domain combination

Centered blurring (i.e. blurring that does not change the centroid²² of the image) acts primarily as a frequency-domain amplitude filter, i.e. a low-pass filter. The phase image is unaffected in a first approximation, however the accuracy of its measurement is determined by the amplitude image [120]. Consequently, views with higher frequency-domain amplitudes are expected to be more reliable than the same frequency-spaces of the alternative views.

The idea behind this group of methods is to fill the frequency space by merging multiple views, whereas each of them is locally (in frequency space) weighted by its local amplitude [133]:

$$\begin{aligned} \tilde{I}(\vec{\omega}) &= \sum_{i=1}^N w_i(\vec{\omega}) \tilde{I}_i(\vec{\omega}) \\ w_i(\vec{\omega}) &= \frac{|\tilde{I}_i(\vec{\omega})|}{\sum_{j=1}^N |\tilde{I}_j(\vec{\omega})|} \end{aligned} \quad (3.47)$$

where $\tilde{I}_i(\vec{\omega}) = \mathcal{F}(I_i(\vec{r}))$ is a Fourier transform of image $I_i(\vec{r})$. A real-domain image is obtained by an inverse Fourier transform $I(\vec{r}) = \mathcal{F}^{-1}(\tilde{I}(\vec{\omega}))$. Some parts in the frequency space will thus be determined predominantly by the views, which cover those parts of the space best, and not by those, that bear weak or no information about the spatial frequencies in question. The algorithm was shown to produce image fusion of four views with nearly isotropic resolution, close to the

²² Centroid, or a center of mass of a gray scale image $I(\vec{r})$ is defined as follows:

$$\vec{r}_c = \frac{\int I(\vec{r}) \vec{r} d\vec{r}}{\int I(\vec{r}) d\vec{r}}$$

Blurring with a PSF will not change the centroid of an image, if and only if the PSF's centroid a null vector.

lateral resolution of input images [133], however no mathematical or physical justification for the procedure was published so far.

3.5.4.7 Maximum-amplitude combination

This method also relies on the fact, that blurring due to a microscope's PSF acts primarily as an amplitude filter (see previous chapter). The image in frequency space can, therefore, be assembled from the multiple views by picking the complex amplitude at every spatial frequency $\vec{\omega}$ from the views with the largest magnitude [164]:

$$\tilde{I}(\vec{\omega}) = \max_i \left(\tilde{I}_i(\vec{\omega}) \right) \quad (3.48)$$

The function max refers to the largest absolute value. Again, a real-domain image is obtained by an inverse Fourier transform: $I(\vec{r}) = \mathcal{F}^{-1}(\tilde{I}(\vec{\omega}))$.

If the blurring function of the microscope is space invariant, it can be described with a point spread function PSF. Due to the convolution theorem, individual views in the frequency domain can be described as a product (3.2). Substituting this into equation (3.48), the following relation is obtained:

$$\begin{aligned} \tilde{I}(\vec{\omega}) &= \max_i \left(\tilde{O}(\vec{\omega}) \cdot \widetilde{PSF}_i(\vec{\omega}) \right) \\ &= \tilde{O}(\vec{\omega}) \cdot \max_i \left(\widetilde{PSF}_i(\vec{\omega}) \right) \\ &= \tilde{O}(\vec{\omega}) \cdot \widetilde{PSF}_e(\vec{\omega}) \end{aligned} \quad (3.49)$$

where \widetilde{PSF}_i represents the Fourier transform of the point spread function of i -th view. Equation (3.49) shows that this fusion method is linear in the sense, that a fused image can be described as a convolution of $O(\vec{r})$ with an effective point spread function $PSF_e(\vec{r})$. The effective PSF is calculated from the PSFs of the individual views by an equation similar to (3.48):

$$\widetilde{PSF}_e(\vec{\omega}) = \max_i \left(\widetilde{PSF}_i(\vec{\omega}) \right). \quad (3.50)$$

This is, therefore, a linear image combination method.

3.5.4.8 Iterated Wiener fusion-deconvolution (MVD-Wiener)

This algorithm following Swoger et al. [90] can simultaneously perform fusion of multiple views and a non-blind deconvolution. The process improves the image estimate iteratively. During every cycle p , a set of simulated images \tilde{S}_j^p generated from the current estimate \tilde{O}^p and the known \widetilde{PSF}_j , are compared to real images \tilde{I}_j . The differences are fed into a regularized inverse filter similar to the Wiener deconvolution [165,166]:

$$\begin{aligned} \tilde{S}_j^p &= \tilde{O}^p \cdot \widetilde{PSF}_j \\ \tilde{c}_j^p &= \frac{\widetilde{PSF}_j^* \cdot (\tilde{I}_j - \tilde{S}_j^p)}{\widetilde{PSF}_j^* \cdot \widetilde{PSF}_j + \mu^2} \end{aligned} \quad (3.51)$$

again, \tilde{f} denotes the Fourier transform of f , \tilde{f}^* is a complex conjugate of \tilde{f} , and μ is a regularization parameter. It is connected to the signal-to-noise level in the images. However, it is normally determined pragmatically [90]. The filter generates a set of additive correction terms, which measure how similar the simulated images and the actually measured images are. The correction terms are used to obtain an improved estimate of the solution:

$$\tilde{O}^{p+1} = \tilde{O}^p + \sum_{j=1}^N w_j \cdot \tilde{c}_j^p \quad (3.52)$$

$$w_j = \frac{\sqrt{|\tilde{I}_j|}}{\sum_j (\sqrt{|\tilde{I}_j|})} \quad (3.53)$$

The corrections suggested by each of the views are additionally weighted by the views' signal-to-noise ratio (SNR). Since the noise is dominated by Poisson statistics, the SNR is estimated to be a square root of the image intensity.

In general, this iterative scheme does not preserve the total intensity and the non-negativity of the estimate. If these are desired characteristics of the output images (as is the case with fluorescence images), the constraints have to be enforced during every iteration of the procedure:

$$\begin{aligned} \text{non-negativity:} \quad & O^p(\vec{r}) \leftarrow \frac{1}{2}(O^p(\vec{r}) + |O^p(\vec{r})|) \\ \text{normalization:} \quad & O^p(\vec{r}) \leftarrow \left(\frac{\int O^1(\vec{r}) d\vec{r}}{\int O^p(\vec{r}) d\vec{r}} \right) \cdot O^p(\vec{r}) \end{aligned} \quad (3.54)$$

In the last equation, the integrated intensity of the estimates are always normalized the initial image integrated image intensity $\int O^1(\vec{r}) d\vec{r}$. The iterative scheme is repeated until it converges to a static solution, i.e. until the integral image difference of an additional iteration is not below a selected threshold ε_{max} : $\int |O^{p+1}(\vec{r}) - O^p(\vec{r})| d\vec{r} > \varepsilon_{max}$. Such solution is thus close to an eigenvector of the every iteration's transformation.

The initial model is created by a weighted average of the input views

$$\tilde{O}^1 = \sum_{j=1}^N w_j \cdot \tilde{I}_j \quad (3.55)$$

where the weights w_j are defined by equation (3.53). This algorithm for a multiple views fusion/deconvolution seems to allow a SPIM to exceed the resolution of confocal microscopy, even when the latter is used with higher NA lenses [90]. However, it has never been shown in a strict mathematical proof that the algorithm increases any kind of merit of the fluorophore distribution estimate. The algorithm also lacks a robust way of determining the optimal value of a regularization parameter μ . The excessively low values make the algorithm inherently unstable while values above an optimal force the algorithm to converge slowly towards a solution that does not take the full advantage of the input views. Last but not least, since some parts of the algorithm

are implemented in frequency space and other parts in real space, a pair of forward and inverse Fourier transforms is required during every iteration of the algorithm. This presents a considerable problem if the input images contain sharp edges (e.g. if fluorescent specimen sticks over the edges of the images). Such images develop “ringing” that spreads from sharp boundaries into the rest of the image, and advances with every additional iteration. The problem can be mitigated by either damping sharp transitions in the input images or by padding the images with a function that slowly decays to the boundaries. Unfortunately, both solutions only seem to replace one challenge with a set of others, e.g. increased processing time and image size.

3.5.4.9 Real-domain based methods

The fusion in the real domain has a number of advantages over algorithms that are implemented in frequency space: i) real-space algorithms can be easily parallelized, ii) they do not suffer from “ringing” artifacts, and iii) they allow the application of space-variant fusion, e.g. quality based, as in section 3.5.4.1, or local SNR based. A number of real space methods are described in the literature and will be sketched in the following sections. In section 3.5.4.14, our own real-domain fusion algorithm is proposed and evaluated.

3.5.4.10 Arithmetic mean

A non-weighted arithmetic mean (or alternatively a sum) of multiple views was proposed several times as a simple and quick way to fuse the views into an image with an improved resolution [133,164]. The method is linear, i.e. its result can be described as a convolution of the fluorophore distribution $O(\vec{r})$ with an effective PSF:

$$\begin{aligned}
 I(\vec{r}) &= \frac{1}{N} \sum_{j=1}^N I_j(\vec{r}) \\
 &= \frac{1}{N} \sum_{j=1}^N O(\vec{r}) * PSF_j(\vec{r}) \\
 &= O(\vec{r}) * \left(\frac{1}{N} \sum_{j=1}^N PSF_j(\vec{r}) \right)
 \end{aligned} \tag{3.56}$$

Here, $*$ denotes convolution. The effective PSF (the expression in the parentheses of the last equation) is an average of the PSFs of the N input views, PSF_j . Due to averaging, such a PSF is obviously more isotropic than any PSF_j of the single input views. Its volume is also considerably decreased (e.g. $2.8\times$ in the case of four SPIM views [133]). However, this method completely disregards the *a priori* information offered by the known PSF_j . The resolution improvement is based solely on the fact that only the central part of all PSF_j is intense in all orientations. If many differently oriented PSFs are averaged, only the central part thus maintains its intensity while the tails “average out”. The technique is, therefore, a bit awkwardly referred to as a “statistical” method [164], although the blur in the images that are rectified is not a statistical artifact.

3.5.4.11 Intensity product

The voxel-wise product of intensities in multiple views, inspired by multiple-photon-like interpretation of multiple observations (section 1.6.3), was suggested as means of fusing multiple images [167]:

$$I(\vec{r}) = \prod_{j=1}^N I_j(\vec{r}) \quad (3.57)$$

Unfortunately, such an interpretation is valid only if photons originating from the same point in the imaged object can be identified in the image [35]. In LSFM this is only possible if the fluorophores are condensed in small clusters. The clusters must be separated well enough, that the fluorescence they emit is detected as clearly separated spots. This method is thus not robust, i.e. it produces legitimate results only if the observed specimen meets certain conditions. It is also not linear, i.e. the fused image can in general (if the condition stated above is not fulfilled) not be described as a convolution of a fluorophore distribution with an effective PSF.

3.5.4.12 Minimum intensity

The PSF of an LSFM spans a longer distance along the axial direction than the lateral direction. If multiple images of an isolated fluorophore or a condensed cluster are recorded along different directions, then only the central part of the spot will be intense in all the views. A point-wise minimum of all the views will, therefore, produce a spot with a reduced size [167]:

$$I(\vec{r}) = \min_j (I_j(\vec{r})) \quad (3.58)$$

Once again, this method is only valid if a specimen with sparsely distributed point-like fluorescence sources is imaged. This fusion algorithm is, therefore, neither robust, nor linear.

3.5.4.13 Iterative constrained Tikhonov restoration from multiple views

Many restoration algorithms [168] solve equation (3.46) by minimizing a functional of a form:

$$\phi = L(\mathbf{H}\mathbf{f}, \mathbf{g}) + \gamma P(\mathbf{f}) \quad (3.59)$$

where $L(\mathbf{H}\mathbf{f}, \mathbf{g})$ is some kind of difference measure between the expected images $\mathbf{H}\mathbf{f}$ and the actually recorded images \mathbf{g} , $P(\mathbf{f})$ is some kind of roughness measure, and γ is a regularization parameter that balances the solution between fitting by L and smoothing by P . Without the last term, the minimization of function (3.59) normally yields solutions for \mathbf{f} that correspond to rugged, highly non-continuous and rather unlikely fluorophore distributions $O(x, y, z)$ (see section 3.5.4.4).

The linear Tikhonov filtering uses the L_2 norm as a measure of both, the fit between $\mathbf{H}\mathbf{f}$ and \mathbf{g} , and the roughness of \mathbf{f} :

$$\phi = \|\mathbf{H}\mathbf{f} - \mathbf{g}\|^2 + \gamma \|\mathbf{C}\mathbf{f}\|^2 \quad (3.60)$$

where \mathbf{C} is a regularization matrix. This is usually a high-pass filter such as a Laplacian, but it can also be an identity matrix in the absence of further information about the expected spectrum of \mathbf{f} [169].

A number of different iterative algorithms for minimizing the functional (3.60) were proposed [168]. One of them, a maximum *a posteriori* with Gaussian noise and prior (MAPGG) was shown to be a feasible method for simultaneous fusion and deconvolution of multiple-view microscope images [90,128,169]. MAPGG applies the transformation $\mathbf{f} = \mathbf{x}^2$ to constrain \mathbf{f} to non-negative solutions $\mathbf{f} \geq 0$:

$$\phi = \|\mathbf{H}\mathbf{x}^2 - \mathbf{g}\|^2 + \gamma\|\mathbf{C}\mathbf{x}^2\|^2 \quad (3.61)$$

To minimize ϕ with respect to \mathbf{x} iteratively, a conjugate gradient algorithm [152] is used. The whole procedure is implemented exclusively in the real-domain.

3.5.4.14 Iterative expectation maximization image fusion

Another algorithm due to W. H. Richardson [170] and L. B. Lucy [163] tackles the inverse problem (3.45) by means of statistical methods. Originally, the problem was stated as the one of estimating a frequency distribution $\psi(\xi)$ of a quantity ξ when it can only be gauged through a finite sample x_1, x_2, \dots, x_n drawn from a population characterized by a distribution

$$\phi(x') = \int \psi(\xi')P(x'|\xi')d\xi' \quad (3.62)$$

Here, $P(x'|\xi')$ is a conditional probability that the measurement of x falls in the interval $(x', x' + dx')$ when it is known that $\xi = \xi'$. This Fredholm integral equation of the first kind is formally equivalent to a linear image creation relation (3.13). Again, if the kernel $P(x'|\xi')$ is a function of $(x' - \xi')$ only, the integral (3.62) becomes a convolution. Note that irrespective of ξ , a randomly chosen x will always take one of the possible values, therefore $\int P(x'|\xi')dx' = 1$.

Let's assume that $Q(\xi'|x')$ is an inverse conditional probability, i.e. a probability that ξ lies in an interval $(\xi', \xi' + d\xi')$ when it is known that $x = x'$ was measured. If a such inverse kernel is known, the distribution $\psi(\xi)$ can be readily obtained by an integral equivalent to (3.62):

$$\psi(\xi') = \int \phi(x')Q(\xi'|x')dx' \quad (3.63)$$

The joint probability that $x \in (x', x' + dx')$ and $\xi \in (\xi', \xi' + d\xi')$ can be expressed in two equivalent ways: $\phi(x')dx' \cdot Q(\xi'|x')d\xi' = \psi(\xi')d\xi' \cdot P(x'|\xi')dx'$. This results in the Bayes' theorem for conditional probabilities:

$$Q(\xi'|x') = \frac{\psi(\xi')P(x'|\xi')}{\phi(x')} = \frac{\psi(\xi')P(x'|\xi')}{\int \psi(\xi')P(x'|\xi')d\xi'} \quad (3.64)$$

Unfortunately, the inverse kernel $Q(\xi'|x')$ in (3.64) is a function of $\psi(\xi')$, which is unknown by definition. Lucy proposed the following iterative procedure to estimate $\psi(\xi')$: form a guess of

$\psi(\xi')$, calculate an estimate inverse kernel (3.64) and then use this kernel to improve the estimate of $\psi(\xi')$. If $\psi^p(\xi')$ is the estimate after p iterations, $\psi^{p+1}(\xi')$ is:

$$\psi^{p+1}(\xi') = \int \phi(x') Q^p(\xi'|x') dx' \quad (3.65)$$

where

$$Q^p(\xi'|x') = \frac{\psi^p(\xi') P(x'|\xi')}{\int \psi^p(\xi'') P(x'|\xi'') d\xi''} \quad (3.66)$$

Such an iterative scheme has the following desirable characteristics [163]:

- It preserves the non-negativity constraint: $\psi^p(\xi') \geq 0 \Rightarrow \psi^{p+1}(\xi') \geq 0$
- It preserves the integral: $\int \psi^{p+1}(\xi') d\xi' = \int \psi^p(\xi') d\xi' = \int \psi^0(\xi') d\xi'$
- $\psi^{p+1}(\xi)$ is more likely to generate measurements $\phi(x')$ than $\psi^p(\xi)$

As Lucy pointed out [163], the procedure is not limited to one-dimensional distributions. In the case of the three dimensional image formation (3.13), the procedure has the following form:

$$M^{p+1}(\vec{r}) = M^p(\vec{r}) \int \frac{I(\vec{q})}{S^p(\vec{q})} h(\vec{r}, \vec{q}) d\vec{q} \quad (3.67)$$

where $I(\vec{q})$ is recorded image, $h(\vec{r}, \vec{q})$ is a blurring function, $M^p(\vec{r})$ is our estimate (model) of the fluorophore distribution after p iterations and $S^p(\vec{q})$ is an image simulated from the estimate $M^p(\vec{r})$:

$$S^p(\vec{q}) = \int M^p(\vec{r}) h(\vec{r}, \vec{q}) d\vec{r} \quad (3.68)$$

Here, \vec{q} is a vector in image space, and \vec{r} is a vector in the space of the model and the object represented by the model. Since $h(\vec{r}, \vec{q})$ is in effect a conditional probability density, its image side (\vec{q} space) integral has to be unity: $\int h(\vec{r}, \vec{q}) d\vec{q} = 1$.

Through those equations, the inner workings of the rectification procedure can be understood intuitively. During every iteration a simulated image (3.68) is calculated according to the current model. The simulated image is then compared to the actual measurement by the division in equation (3.67). Fraction $\frac{I(\vec{q})}{S^p(\vec{q})}$ is greater than unity where the simulated image is darker than the recorded image, and less than unity where the simulated image is brighter. By multiplying this ratio with the model, an improved model is obtained. However, if the fraction indicates a mismatch in a point \vec{q} , this might be due to an error in the model anywhere in the volume \vec{r} , where $h(\vec{r}, \vec{q}) \neq 0$. Convolving the division with the blurring function $h(\vec{r}, \vec{q})$ in equation (3.67) ensures that the change is properly distributed and weighted by the influence that different regions of the model have on the locations of mismatch. Such weighting seems a sensible approach in the absence of additional information, even without the stringent validation provided by the mathematical derivation above.

Following equation (3.67) it is obvious that the low-frequency discrepancies (i.e. the deviations on length scales that are large compared to the scale of $h(\vec{r}, \vec{q})$) between the simulated image and the recorded image will be corrected effectively in a single iteration. The choice of an initial model $M^0(\vec{r})$ is, therefore, not very important, as long as it is reasonably smooth and has an integral intensity expected from the final image. On the other hand, the deviations on the length scale of $h(\vec{r}, \vec{q})$ and shorter cause relatively small corrections to the model. This prevents high-frequency noise from being excessively amplified.

Extension to multiple-view

We extended the algorithm to the multiple-view image reconstruction problem. The kernel function $P(x'|\xi')$ in the integral (3.62) can be extended to connect the distributions in two spaces of arbitrary dimensionality [163]. The multiple-view image formation process can be described by a formula similar to a standard linear image formation model:

$$I(\vec{q}, i) = \int O(\vec{r}) h(\vec{r}, \vec{q}, i) d\vec{r} \quad (3.69)$$

Here, $i = 1 \dots N$ represents the views along different directions and $I(\vec{q}, i)$ is now a four-dimensional (three continuous and one discrete dimension) image set: $I(\vec{q}, i) = I_i(\vec{q})$, where $I_i(\vec{q})$ denotes a single image acquired along direction i . Again, \vec{q} is a vector in the image space of individual three-dimensional images $I_i(\vec{q})$ and \vec{r} is a vector in the space of the object. The integral kernel $h(\vec{r}, \vec{q}, i)$ now links a three-dimensional object space of \vec{r} and a four dimensional image space (\vec{q}, i) and can be described as a set of blurring functions of the individual views:

$$h(\vec{r}, \vec{q}, i) = \frac{1}{N} \cdot h_i(\vec{r}, \vec{q}) \quad (3.70)$$

where $h_i(\vec{r}, \vec{q})$ is a blurring function of a view i . The normalization factor $1/N$ is derived from the probabilistic nature of the integration kernels; their image-side (i.e. \vec{q} space) integrals are equal to unity (see page 108):

$$\int h_i(\vec{r}, \vec{q}) d\vec{q} = 1$$

$$\sum_{i=1}^N \int h(\vec{r}, \vec{q}, i) d\vec{q} = 1$$

The reconstruction of a fluorophore distribution model $M(\vec{r})$ based on the information from multiple views $I_i(\vec{q})$ is thus described by the following iterative scheme:

$$M^{p+1}(\vec{r}) = M^p(\vec{r}) \sum_{i=1}^N \int \frac{I(\vec{q}, i)}{S^p(\vec{q}, i)} h(\vec{r}, \vec{q}, i) d\vec{q} \quad (3.71)$$

$$S^p(\vec{q}, i) = \int M^p(\vec{r}) h(\vec{r}, \vec{q}, i) d\vec{r}$$

Equivalently, equations (3.71) can be expressed exclusively by three dimensional images $I_i(\vec{q})$ and a set of standard blurring functions (3.70), $h_i(\vec{r}, \vec{q})$:

$$M^{p+1}(\vec{r}) = M^p(\vec{r}) \frac{1}{N} \sum_{i=1}^N C_i(\vec{r}) = M^p(\vec{r}) \cdot \langle C_i^p(\vec{r}) \rangle \quad (3.72)$$

$$C_i^p(\vec{r}) = \int \frac{I_i(\vec{q})}{S_i^p(\vec{q})} h_i(\vec{r}, \vec{q}) d\vec{q} \quad (3.73)$$

$$S_i^p(\vec{q}) = \int M^p(\vec{r}) h_i(\vec{r}, \vec{q}) d\vec{r} \quad (3.74)$$

Here, $C_i(\vec{r})$ is a correction to the model $M(\vec{r})$ suggested by the view i . The first equation reveals that the final correction to the model is simply an average of the corrections suggested by the individual views. This is of utmost importance for an efficient computational implementation of the procedure. The individual images $I_i(\vec{q})$ can be compared to simulated images $S_i^p(\vec{q})$ independently for every view i . Only one image and its simulated counterpart must, therefore, be simultaneously contained in the computer's memory. Once all correction functions $C_i(\vec{r})$ are calculated the information in all views is fused simply as an average $\langle C_i(\vec{r}) \rangle$.

The formulation can be further simplified if $h_i(\vec{r}, \vec{q})$ are functions of $(\vec{q} - \vec{r})$ ²³: $h_i(\vec{r}, \vec{q}) = PSF_i(\vec{q} - \vec{r})$. The formation of N views can then be expressed as N convolutions:

$$\begin{aligned} I_1(\vec{q}) &= \int O(\vec{r}) PSF_1(\vec{q} - \vec{r}) d\vec{r} \\ &= O(\vec{q}) * PSF_1(\vec{q}) \\ I_2(\vec{q}) &= O(\vec{q}) * PSF_2(\vec{q}) \\ &\vdots \\ I_N(\vec{q}) &= O(\vec{q}) * PSF_N(\vec{q}) \end{aligned} \quad (3.75)$$

The rectification procedure is then expressed as follows:

$$M^{p+1}(\vec{q}) = M^p(\vec{q}) \cdot \langle C_i(\vec{q}) \rangle \quad (3.76)$$

$$\langle C_i(\vec{q}) \rangle = \frac{1}{N} \sum_{i=1}^N \left(\left(\frac{I_i(\vec{q})}{M^p(\vec{q}) * PSF_i(\vec{q})} \right) * PSF_i'(\vec{q}) \right) \quad (3.77)$$

where PSF_i' is a space-inverted PSF_i : $PSF_i'(\vec{q}) = PSF_i(-\vec{q})$.²⁴

Again, the choice of the initial model $M^0(\vec{r})$ is not important as long as it is a smooth function of \vec{r} and it has an integral intensity $\int M^0(\vec{r}) d\vec{r}$ expected from the final result. An average of all input views seems a reasonable choice:

²³ Please note that this implies that \vec{r} and \vec{q} are vectors in the same space. The distinction between vectors in image space and vectors in object/model space is thus not required anymore. This applies to all ensuing equations.

²⁴ This is due to the fact, that in equation (3.73) the integration is performed with respect to \vec{q} while in equation (3.74), it is performed with respect to \vec{r} . However, most microscopes' point spread functions can be approximated by a symmetric PSF: $PSF_i'(\vec{q}) = PSF_i(\vec{q})$.

$$M^0(\vec{q}) = \frac{1}{N} \sum_{i=1}^N I_i(\vec{q}) \quad (3.78)$$

Numerical evaluation

The single image restoration defined by the iterative scheme (3.67) and (3.68) is usually referred to as Lucy-Richardson deconvolution. It rectifies an image that was distorted by a low-pass amplitude filter (i.e. blurred), by restoring the high frequencies that are still present in the image but are attenuated. Figure 53 shows how the Lucy-Richardson deconvolution inflates the spectrum of simulated image.

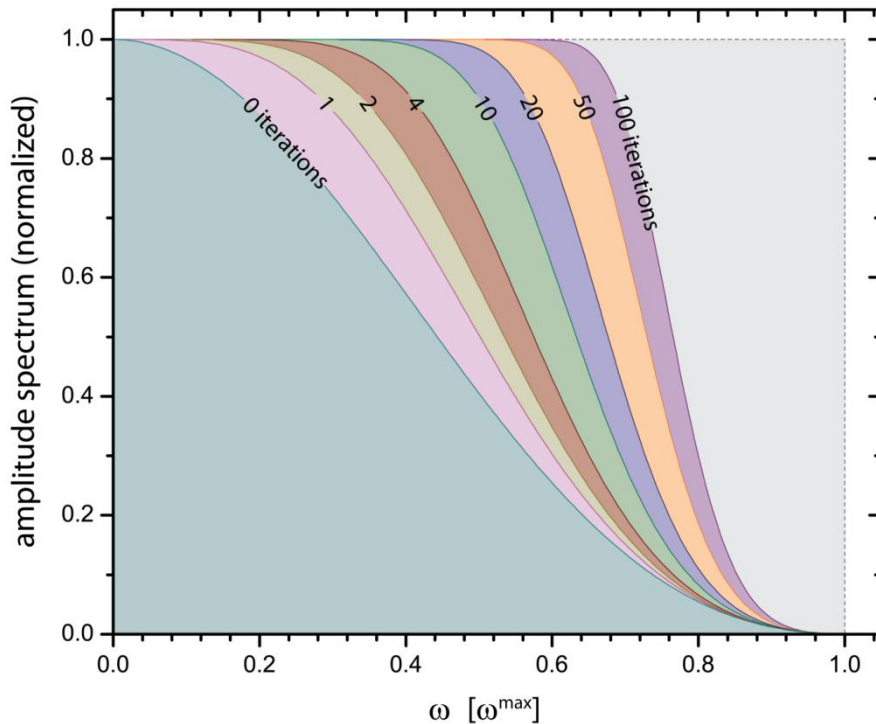


Figure 53: Image amplitude spectrum envelope and Lucy-Richardson deconvolution. The imaging process behaves like a low pass filter, which demonstrates in the bounded OTF. The high frequencies in the image are progressively attenuated up to the cut-off frequency (ω^{max}) above which the amplitude envelope is zero. The deconvolution algorithms amplify the frequencies that are attenuated during the imaging process. The plot, resulting from a numerical simulation, shows how an image's spectral envelope expands with an increasing number of Lucy-Richardson deconvolution iterations (see text for details). The increase is most rapid for the low and central frequencies of the spectrum, while the high frequencies are recovered much less. The frequencies above the cut-off ω^{max} cannot be recovered. The gray rectangle indicates the theoretically attainable spectrum, where all frequencies contained in the original image have equal amplitudes/contrast. In reality, this level of restoration is impossible due to noise that drowns the highly attenuated high frequencies near the cut-off. Abscissa is normalized to the ω^{max} and ordinate is normalized to the highest amplitude. Graphs were produced by a simulated deconvolution of point-object image.

Images consisting of isolated point objects that were convolved with a synthetic PSF kernel. The PSF was chosen such that its spectrum was limited to frequencies below a cut-off frequency ω^{max} :

$$MTF(\omega) = |\widetilde{PSF}(\omega)| = \begin{cases} \text{sinc}\left(\pi \frac{|\omega|}{\omega^{max}}\right) & \text{if } |\omega| < \omega^{max} \\ 0 & \text{otherwise} \end{cases} \quad (3.79)$$

This particular function was chosen because it closely resembles the modulation transfer functions of real microscopes, it has a bounded support and it is reasonably smooth²⁵. The blurred (one-dimensional) images were then rectified using the Lucy-Richardson algorithm. Without deconvolution, the spectrum of an image is limited by an envelope defined by the shape of the amplitude filter (3.79).

As demonstrated by Figure 53, the deconvolution gradually extends the envelope, which covers an increasing fraction of the theoretically attainable spectrum. The attainable spectrum is fundamentally limited to a band $|\omega| < \omega^{max}$ since no higher frequencies remain in the image after it was filtered by (3.79). Figure shows that the image energy $\int MTF(\omega)d\omega$, as a share of the total attainable spectrum (gray rectangle), increased from 45% to 77% within the first 100 iterations. The gain was mainly in the middle-frequency part of the spectrum, while the amplification of the frequencies close to the cut-off frequency ω^{max} is more than modest. As already mentioned above, this is a desirable characteristic of a deconvolution algorithm; in the case of real microscopy images, the highly attenuated image frequencies near the cut-off ω^{max} are inevitably drowned in noise and should thus not be amplified.

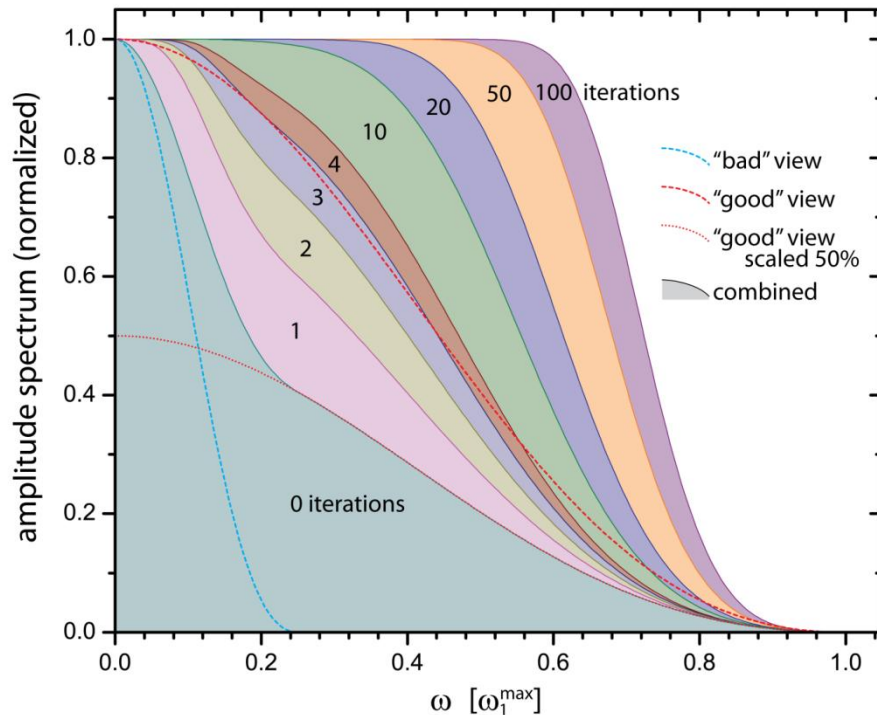


Figure 54: Image amplitude spectrum envelope and iterative expectation maximization fusion algorithm.

The plot shows how the amplitude spectrum envelope of an image expands with an increasing number of fusion iterations. In this example, a “good” and a “bad” view with cut-off frequencies $\omega_1^{max} = 5 \cdot \omega_2^{max}$ (red and blue dashed lines, respectively) are fused. The initial model (0 iterations) is a mean of the “good” and the “bad” views. Its spectrum envelope is a mean of respective spectrum envelopes whose high frequencies are contributed solely by the “good” view (red dotted line is a 50% scaled “good” view). The spectral envelope is quickly recovered and within 5 iterations, it covers the same area as the “good” view alone. Additional iterations expand the spectrum further. From this point on, the fusion’s spectrum envelope is defined mostly by the spectrum envelope of the “good” view (compare with Figure 53). Abscissa is normalized to the ω_1^{max} and ordinate is normalized to the highest amplitude.

²⁵Different functions with such properties were evaluated without significant effects on the outcome of the simulation.

Within the one-dimensional analysis framework described above, a multiple-view set corresponds to a collection of one-dimensional measurements with different MTFs and different cut-off frequencies ω_i^{max} , where $i = 1 \dots N$ are different views. Such a set of measurements is equivalent to a set of one-dimensional image profiles along an arbitrary direction in a multiple-view image set. The fusion is expected to take full advantage of the information contained in the “best” view (i.e. the view with the highest cut-off frequency along the chosen direction) even in the presence of “bad” views. In the one-dimensional case this sounds like a trivial task: “bad” views can be easily identified and simply ignored, thus avoiding a multiple-view fusion. However, in the case of three-dimensional images, every view has one “bad” direction and many “good” directions. The fusion cannot, therefore, rely on one view alone, but must combine “good” information from all of them.

The simulation of fused images of a point object, which were blurred by a set of synthetic, bounded PSFs corresponds to the MTFs from equation (3.79) with different ω_i^{max} . For the sake of simplicity, the first view was chosen to have the highest cut-off frequency, while the other view were considerably more limited (4-6 fold, as in real multiple-view microscopy). Figure 54 shows the results of such a simulation, based on the fusion of two measurements ($\omega_1^{max} = 4 \cdot \omega_2^{max}$; similar to LSFM’s PSF extensions). The initial model was a mean of both measurements. Its spectrum covered only around 28% of the spectrum attainable by the “best”, i.e. the first view; a rather poor start considering that the first view alone covers more than 45% of the spectrum. The shape of the initial model’s spectrum matches the sum of the spectra of the individual views (see dotted line in Figure 54). However, only after 4 iterations, the model spectrum’s shape closely resembles that of the “good” view. From there on, the spectrum is gradually inflated as in the case of a single view (Figure 53). The procedure, therefore, simultaneously fuses and deconvolves the data.

Convergence speed

The convergence speed of the iterative algorithm above was tested using the coverage of the attainable amplitude spectrum as a measure of its effectiveness. The simulation was carried out using the same one-dimensional spectrum analysis as described above. One “good” view with cut-off frequency of ω_1^{max} was fused with a different number of “bad” views: with $\omega_2^{max} = \frac{1}{4}\omega_1^{max}$ in the case of a 2-view fusion and with $\omega_2^{max} = \frac{1}{4}\omega_1^{max}$, $\omega_3^{max} = \frac{1}{5}\omega_1^{max}$ and $\omega_4^{max} = \frac{3}{20}\omega_1^{max}$ in case of a 4-view fusion.

The restoration effectiveness after a number of iterations is defined by the algorithm’s efficiency and by the aptness of the initial model. To isolate the influence of the former, the initial model was always initialized to the same function, i.e. the “good” view. The attainable spectrum coverage in case of 1-, 2- and 4-view based restoration is shown in Figure 55 (solid lines). After approx. 10 iterations, all three curves run in parallel, forming linearly increasing functions of the iteration number on a semi-logarithmic plot. The convergence is, therefore, logarithmic, with its speed decreasing as it approaches the stable solution.

It is clear from the Figure 55 that the convergence speed is also reduced with an increasing number of “bad” views. The fact that all three curves (1-view, 2-view and 4-view fusion) follow parallel lines on the semi-logarithmic plot indicates that the ratios between the convergence

speeds in the three situations are constant. Figure 56 shows how many iterations of the “good”-view deconvolution (black line) yields equivalent results as a given number of iterations of a 2-view (blue line) and 4-view (red line) fusion²⁶. The plots are clearly straight lines with slopes of 2 and 4 in case of 2-view and 4-view fusion, respectively. This shows that the convergence speed decreases linearly with the number of views. This could be attributed to the fact, that every additional view increases the effective volume of the four-dimensional blurring function $h(\vec{r}, \vec{q}, i)$ in equation (3.71). Each point in the final model is influenced by number of image voxels proportional to the number of views. If additional views add no useful information about that location it will only slow down the convergence by decreasing the correction to the model suggested by better views. The final result, however, is not affected.

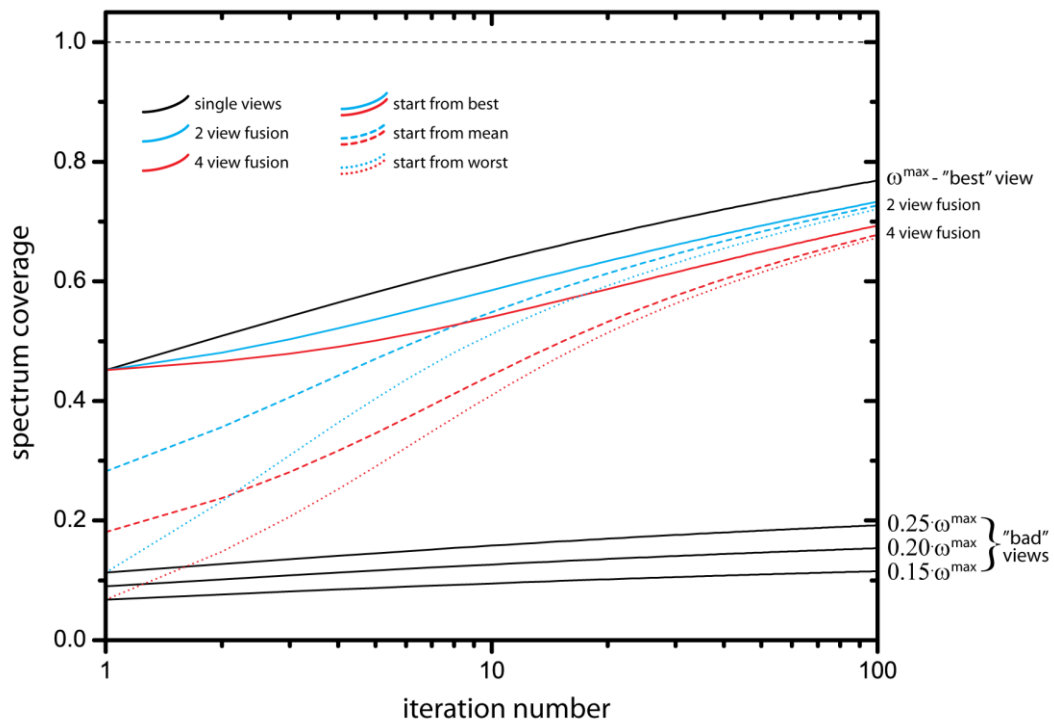


Figure 55: Attainable spectrum coverage as a function of number of fusion iterations. The fusion’s spectrum can contain only spatial frequencies that are included in at least one of the input views. The part of the spectrum that is attainable by a fusion is, therefore, fundamentally limited (e.g. see Figure 54). The plot shows what part of the attainable spectrum is covered with an increasing number of fusion iterations. The results of simulated 2-view (blue lines) and 4-view (red lines) fusions are shown together with the results of a deconvolution of the “best” view (black line). The simulation fused one “good” view and a number (1 or 3) of “bad” views. The three possible initial models are explored: fusion starting from the “best” view (solid line), from the “worst” view (dotted line) and from the mean (dashed line). Within 100 iterations, very similar shares of the spectrum are covered, regardless of the initial model. Within 10 iterations, the solid lines appear to become parallel lines. This indicates that the spectrum coverage is a logarithmic function of the number of iterations. Furthermore, the constant distance between the lines shows that the increasing number of views slows the convergence by a constant factor (see text).

²⁶ Again, only one view spans a wide spectrum while the others contribute only low-frequency data. In such a one-dimensional case, the algorithm is expected to base its estimate primarily on the one “good” view, disregarding the others. As a consequence, the multiple-view fusion of one-dimensional data will always yield poorer results and will converge slower than the “good” view alone. But, one-dimensional data is only used to evaluate the algorithm. In the case of three-dimensional images every view contributes high frequencies in a different part of the frequency/phase space and, therefore, provides useful information that improves the final result and algorithm’s convergence.

The convergence speed of the iterative fusion/deconvolution scheme can be increased by exponentiating the correction factor $\langle C_i(\vec{q}) \rangle$ in (3.76) by a factor $P(p)$:

$$M^{p+1}(\vec{q}) = M^p(\vec{q}) \cdot \langle C_i(\vec{q}) \rangle^{P(p)} \quad (3.80)$$

Values $P(p) > 1$ usually improve the convergence, but they might also lead to an unstable evaluation. It has been reported [171] that single image deconvolution only converges if $P(p) < 3$. An alternative approach is to adaptively increase the value of $P(p)$ to boost the convergence as the stable solution is approached. One of the recent implementations of this idea, an adaptively accelerated Lucy-Richardson (AALR) method [171], suggests calculating the optimal $P(p)$ from first-order derivatives (similarly as discussed in section 97) of $M^{p-1}(\vec{q})$ and $M^{p-2}(\vec{q})$:

$$P(p) = \exp\left(\frac{\|\nabla M^{p-1}(\vec{q})\|}{\|\nabla M^{p-2}(\vec{q})\|}\right) - \left(\frac{\|\nabla M^1(\vec{q})\|}{\|\nabla M^0(\vec{q})\|}\right) \quad (3.81)$$

where \exp is an exponential function, ∇ is first order derivative operator and $\|\cdot\|$ is a L_2 norm. An adaptive acceleration is reported to reduce the number of required iterations for approx. 43%. This method was not tested on the fusion of multiple views yet, but it seems a feasible way to speed-up the fusion algorithm in the case of time-critical applications.

Choice of the initial model

In the previous section, all restoration procedures started with the “best” view. This was possible since simulations were carried out with one-dimensional data. Every real three-dimensional image contains a “bad” direction (i.e. direction with spectrum limited to low frequencies) orthogonal to the plane of “good” directions. “Good” and “bad” data is thus fundamentally interwoven and cannot be separated from a single view alone. The iterative multiple-view fusion algorithm, therefore, cannot start from the “best” view but rather from a model, that treats all views equally.

The effects of the initial model can be seen in Figure 56. Unsurprisingly, the best results (i.e. the highest image energy) are obtained when the initial model is the “best” view (solid lines) and the worst results follow from the “worst” view (dotted lines). However, as already mentioned, this is only feasible with one-dimensional data. When the average of all input views (3.78) is chosen as a start, the spectrum coverage lies somewhere between the two extremes (dashed lines). The models also converge relatively quickly. If the mean of the views is chosen as an initial model, the spectrum coverage is only 0.8% and 2.2% worse (for two- and four-view fusion, respectively) after 100 iterations, than it is if the restoration starts from the “best” view.

Figure 56 also reveals that the convergence speed is not significantly affected by the choice of the initial model. The slope of the convergence depends primarily on the number of views. The suboptimal choice of initial model seems to increase the number of required iterations by a relatively small number of iterations (Figure 56, inset). In the case of two views (blue lines), 3 additional iterations are required to improve the mean initial model to the level of the “best” view and 5 iterations are required for the “worst” view. In the case of four views, the figures are 10 and 12 iterations for mean and “worst” view, respectively. However, the most important result is that the influence of the initial model is relatively small, compared to the total number of iterations required for rectification (~ 100). The mean of all views (3.78) is therefore a reasonable choice.

Fusion of real point-object images

One-dimensional images of a point-object were used in the preceding simulations. Similar results are obtained by fusing three-dimensional images of point objects, i.e. using data recorded with a SPIM. The tiny fluorescent beads are most often used in microscope evaluation as sub-wavelength sources of fluorescence. A large number of 100nm fluorescent beads was dispersed in agarose gel and imaged with an LSFM and a $40\times/0.8\text{W}$ objective lens. Resolutions of approximately $0.36\mu\text{m}$ and $1.6\mu\text{m}$ are expected along the lateral and axial directions, respectively (see section 2.2.3.3). The beads with a 100nm diameter, therefore, have a size that is well below the resolution limit. Their extents in the images are governed primarily by the LSFM's PSF.

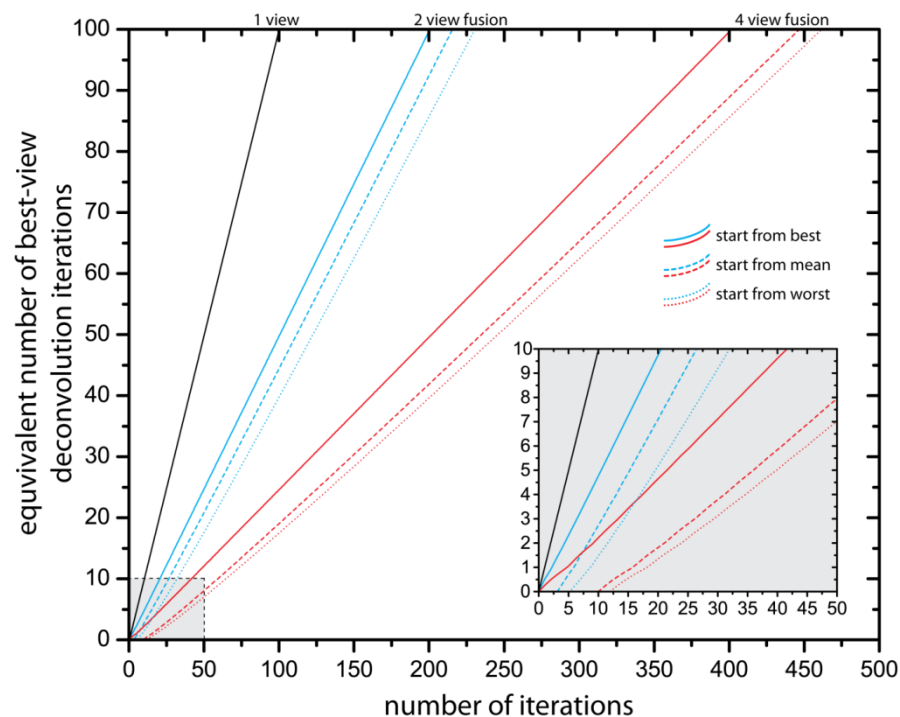


Figure 56: Convergence speed as a function of the number of views. The plot shows how the restoration speed (measured as the attainable spectrum coverage after certain number of iterations, see Figure 55) of multiple views fusion compares to the single-view deconvolution of the “best” view alone. The slopes of the lines are proportional to the number of the fused views, which indicates that the fusion of more views takes longer to achieve equivalent results. The initial model has only limited effect on the slope, but it delays the onset of the restoration (see inset).

Two views along two orthogonal directions were acquired, aligned by the affine transformation (page 87) and fused using the iterative expectation maximization method. The results can be seen in Figure 57. Even the fusion of only two views considerably improves the resolution and makes it almost isotropic. The frequency space is filled accordingly, as suggested by the one-dimensional simulation (Figure 57d-f).

The decrease in the spot size due to multiple-view image fusion/deconvolution is demonstrated by Figure 58. One of the beads from Figure 57 was selected and the extents of its image were measured as the fusion/deconvolution progressed. The spots size was measured by fitting its profile with a Gaussian function. As rotation axis in an LSFM is always normal to the detection optical axis (section 2.2), the spot extension along the rotation axis is always defined solely by the

lateral resolution of the LSFM. The spots always have the smallest extent along the rotation axis. On the other hand, the spots are elongated along a direction normal to the rotation axis.

Figure 58 shows how the spots full width at half maximum along its largest extent (blue solid line) and its smallest extent (red line) decreases during the first 50 fusion/deconvolution iterations. Its smallest extent is reduced from 517nm to 300nm (i.e. approx. $1.7\times$ less) due to the deconvolution effect of the algorithm. The spot's largest extent, on the other hand, was reduced from $1.3\mu\text{m}$ to 323nm (i.e. $4\times$ less), driven primarily by the two-view fusion. The ratio between the smallest and the largest spot extent is reduced from 2.5 to about 1.07; the resulting image appears significantly more isotropic (Figure 58c).

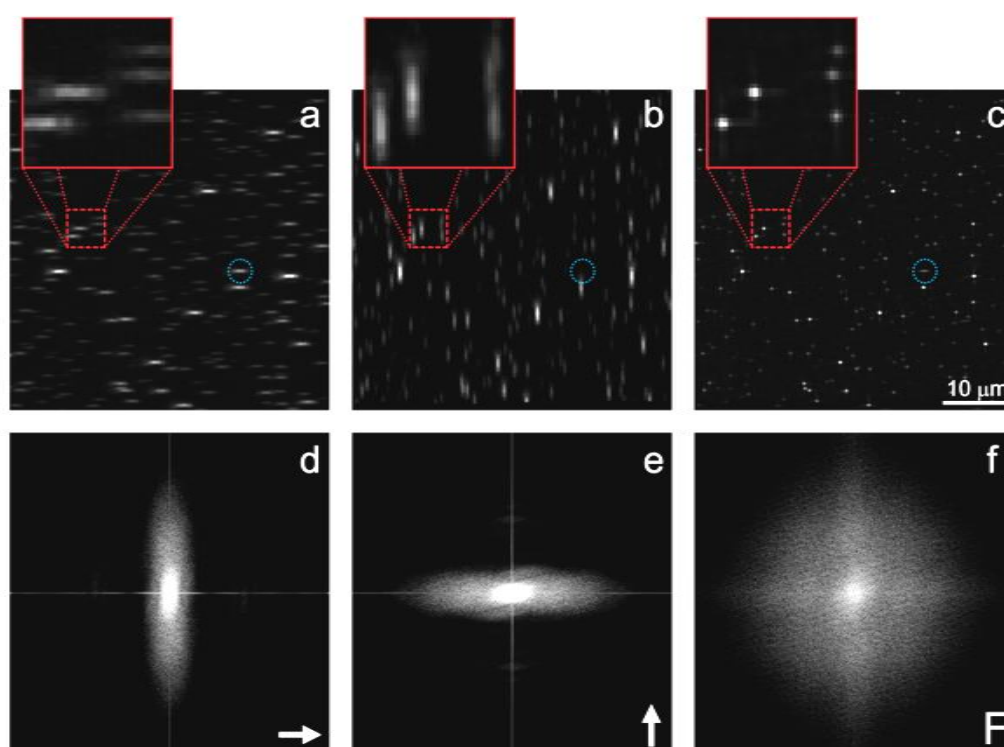


Figure 57: Fusion of images of point-object. **a, b** The fluorescent beads (100nm in diameter) were recorded along two orthogonal directions. **c** shows a fusion of both images. **d-f** show the amplitude spectra of the respective frequency space representations. The lateral and axial point spread function extents in **a** and **b** are $0.4\mu\text{m}$ and $1.3\mu\text{m}$, respectively. The axial extent of the point spread function in **f** is thus improved by a factor of four (4). The arrows in **d** and **e** refer to the directions of the respective detection axes. The real space images are maximum intensity projections. The original data set has a size of $1024\times 1344\times 512$ picture elements. The images were recorded with a Zeiss Achroplan $40\times/0.8\text{W}$, the excitation wavelength was $0.488\mu\text{m}$ and the emission was recorded above $0.510\mu\text{m}$. The scale bar in **c** is valid for **a-c**. Blue circles in **a-c** mark position of a mobile bead, that was recorded in one view only (**a**) and therefore produced an anisotropic, elongated image in the fusion.

Resolution improvement

Resolution is traditionally defined as the minimal distance between two point objects that discriminates them with a well-defined contrast (Cf. Rayleigh and Sparrow criterion [31]). The iterative expectation-maximization image fusion/deconvolution algorithm is not linear in the sense, that the final image can be described as a convolution of an underlying fluorophore distribution with an effective PSF. The resolution of a fused image can, therefore, not be determined simply from the extent of a point-object's image.

The resolution improvement was determined from a series of simulated multiple-view images. The images were generated by emulating LSMF imaging of two point-objects with a varying distance between them. The resolution was then determined according to the Sparrow criterion, i.e. as the minimal distance between the point-objects that allows their discrimination with a non-zero contrast. The resolution was determined for a different number of views and iterations.

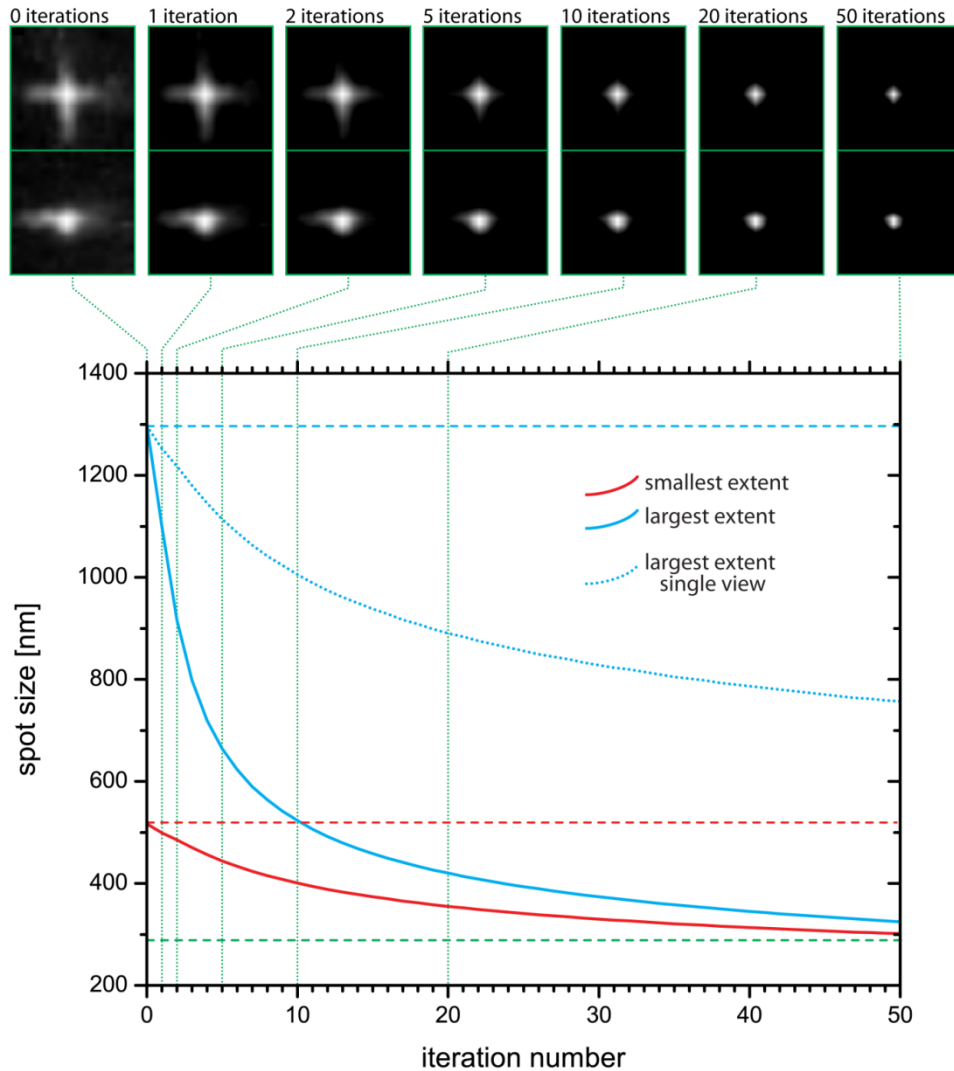


Figure 58: Size of a point-object after a number of two-view fusion iterations. A fluorescent bead from Figure 57 was selected and its extents were measured as the fusion progressed. The plot shows how bead's largest (blue solid line) and smallest (red solid line) extents decreased in the course of the process. The bead's smallest extent decreases solely due to the deconvolution capability of the algorithm, while its largest extent is driven primarily by the fusion of the two views; if only a single view is processed, the beads largest extent stays considerably large (blue dotted line). Within 50 iterations of two-view fusion, the difference between the two extents is reduced from 2.5× to only about 7%. The green line indicates the projected limit of the both functions. The images of the bead for various numbers of iterations are shown above the plot. The upper row is a view parallel to the rotation axis and the bottom row is a view along the detection axis of one of the views.

The PSF was approximated by a three-dimensional Gaussian:

$$G(x', y', z') = \frac{1}{\sigma_t^2 \sigma_a (2\pi)^{\frac{3}{2}}} \exp\left(-\frac{1}{2}\left(\frac{x'^2}{\sigma_a^2} + \frac{y'^2}{\sigma_t^2} + \frac{z'^2}{\sigma_t^2}\right)\right) \quad (3.82)$$

Here, σ_l defines the lateral extent of the PSF, $\sigma_a > \sigma_l$ defines axial extent of the PSF and (x', y', z') is a vector in a coordinate system, which is aligned with the optical system of the LSFM, so that x' is parallel to the detection optical axes:

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} \cos \gamma & -\sin \gamma & 0 \\ \sin \gamma & \cos \gamma & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad (3.83)$$

where, (x, y, z) are image coordinates, γ defines the orientation of the PSF in the image and the axis of rotation is assumed to be parallel to $z = z'$. The N view directions (i.e. PSF orientations) were evenly distributed from 0 to π (opposing views don't contribute any additional information in transparent specimens), so that $\gamma_i = (i - 1)\frac{\pi}{N}$ for different views $i = 1 \dots N$. When a finite number of views is fused, the resolution is not isotropic and depends on the direction. In general, a resolution in an arbitrary direction always lies somewhere between the best resolution and the worst resolution. The best resolution is measured along a direction that is a lateral direction of any of the input views.

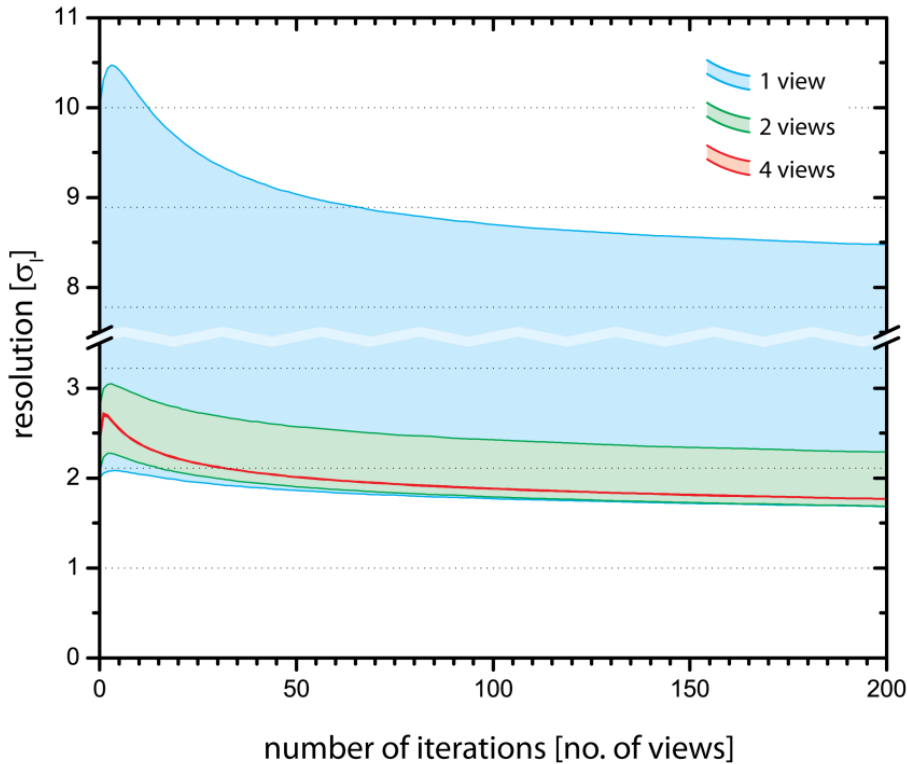


Figure 59: Resolution improvement due to fusion/deconvolution. The resolution according to Sparrow's criterion was evaluated by simulating multiple-view imaging of two point objects. In general, the resolution depends on the orientation of the line that connects both objects relative to the detection optical axis. The resolution is the worst along the optical axis (axial resolution) and best orthogonally to it (lateral resolution; section 2.2.3.3). The resolution along an arbitrary direction will always lie somewhere in between those two extremes. The plot shows how the range of resolutions shrinks as more views are fused. The PSF was assumed to be 5× longer than it is wide. The resolution in this case becomes nearly isotropic when at least 4 views are fused and is only marginally worse than the lateral resolution of the single views. The number of iterations was scaled with the number of views to balance the decrease in convergence speed (see page 114); e.g. in case of 2-view and 4-view fusion, the abscise range is 0-400 iterations and 0-800 iterations, respectively.

Both, the best and the worst resolution were measured in this simulation. The PSF was chosen to be elongated similarly as in LSMF (section 2.2.3.3): $\sigma_a = 5 \cdot \sigma_l$. The number of iterations was scaled with the number of fused views to compensate for the decrease in convergence speed. Results are demonstrated in Figure 59.

Additional views clearly improve the worst resolution in an anisotropic image. The gain is biggest when the second view is added. Four views produce an image with approximately isotropic resolution. If the PSF was more elongated (it rarely is in LSMF), more views would be required to reach this level of isotropy. Interestingly, the resolution initially degrades (for approx. 5-10%) with the first 4 iterations, before it starts to improve.

Figure 60 shows the resolution improvement as a function of a number of views. The resolution simulation was carried out using the same presumptions as for Figure 59. Again, the resolution improves most with the first additional view, while there is practically no gain beyond 4 views (in the case of 5-fold PSF elongation).

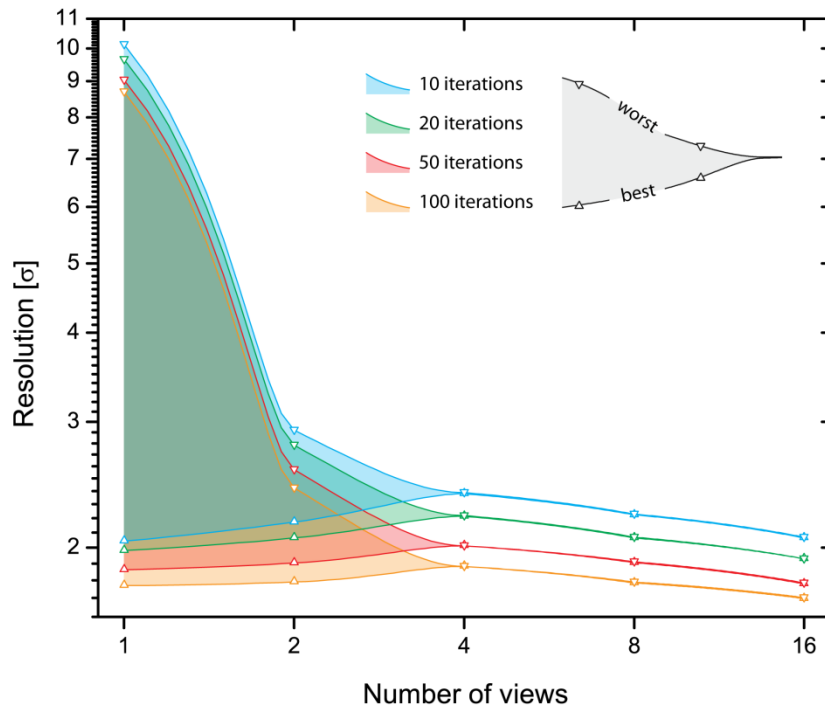


Figure 60: Resolution improvement due to fusion/deconvolution. The plot shows data from Figure 59 as a function of the number of views. Again, resolution becomes isotropic when at least four views are fused. Interestingly, the “best” resolution is first increased for approx. 5% as it is approached by the “worst” resolution, but it is improved again beyond the original lateral resolution, if more than four views are fused.

Fusion of simulated multiple-view images

The algorithm was further tested on a series of anisotropically blurred images. The source image was acquired by SPIM, and blurred by a set of kernels (3.82) with $\sigma_a = 10$ pixels and $\sigma_l = 2$ pixels. Again, the N view directions (i.e. the PSF orientations) were evenly distributed from 0 to π . The source image and the four simulated views are shown in Figure 61. The orientation of the blur is clearly observable in each of them; the patterns in the image that have a similar orientation as the blur are amplified, while the patterns with dissimilar orientations are attenuated. The deconvolution of individual views (Figure 62, left column) improves the sharpness of the images, but the remaining blur is still highly anisotropic. For example, the bottom-left image in Figure 62

shows that even after 50 deconvolution iterations of an image that was blurred with a vertically oriented kernel, the horizontal lines are still not discernable (see the zoomed region). They become visible when the second, vertical view is added.

Again, the improvement is most clearly visible when the number of views is increased to four. After that, the improvement by doubling the number of views becomes hardly perceptible. However, this only applies to the blur settings applied in the simulations ($\sigma_a = 5 \cdot \sigma_l$). A more elongated kernel requires more views.

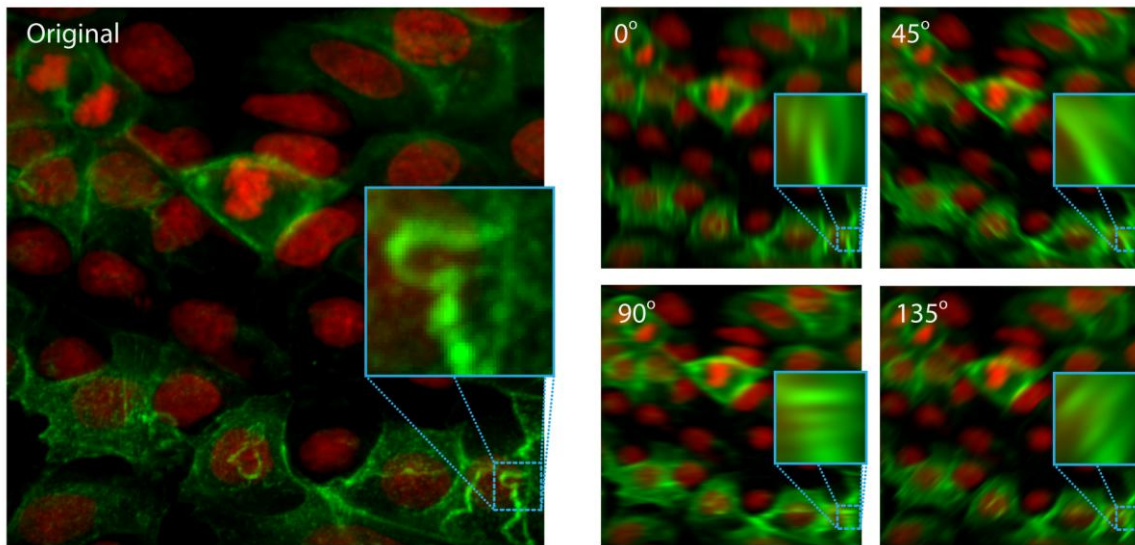


Figure 61: Fusion of simulated multiple-view images. The original image (left) and four blurred images, simulating multiple-view imaging that are used as an input for the multiple-view image fusion. PSF was approximated by a three dimensional Gaussian, that is $5\times$ as long as it was wide. The angle values in the four images on the right denote the direction of the PSF's long axis in the images (0° is vertical). The results of the fusion are show in Figure 62. The original image shows MDCK cells cultured on a glass cover slip. Actin (shown in green) was stained by Alexa-488 Phalloidin (Mocleular Probes) and nuclei (shown in red) by Draq5 (Biostatus). Cells were image by SPIM and Carl Zeiss Achroplan 40x/0.75W objective lens; illumination at 488nm and 647nm and detection above the respective wavelengths (long-pass filter). Length of the image's edge corresponds to $25\mu\text{m}$.

The effectiveness of the image restoration can be evaluated by comparing the output of the algorithm with the source image that is rectified in the simulated views. The normalized cross-correlation (3.20) was used as a measure of how close to the source image the estimates came. Figure 63 shows how an image difference function $(1 - NCC(I_1, I_2))$ decreases with an increasing number of algorithm iterations and/or input views. Again, the improvement is most significant when number of views is increased from 1 to 4. Apparently, adding another 4 views (8 in total) improves the image a bit further, but it hardly justifies having to acquire double the number of images. Improvement provided by fusing more than 8 views is negligible.

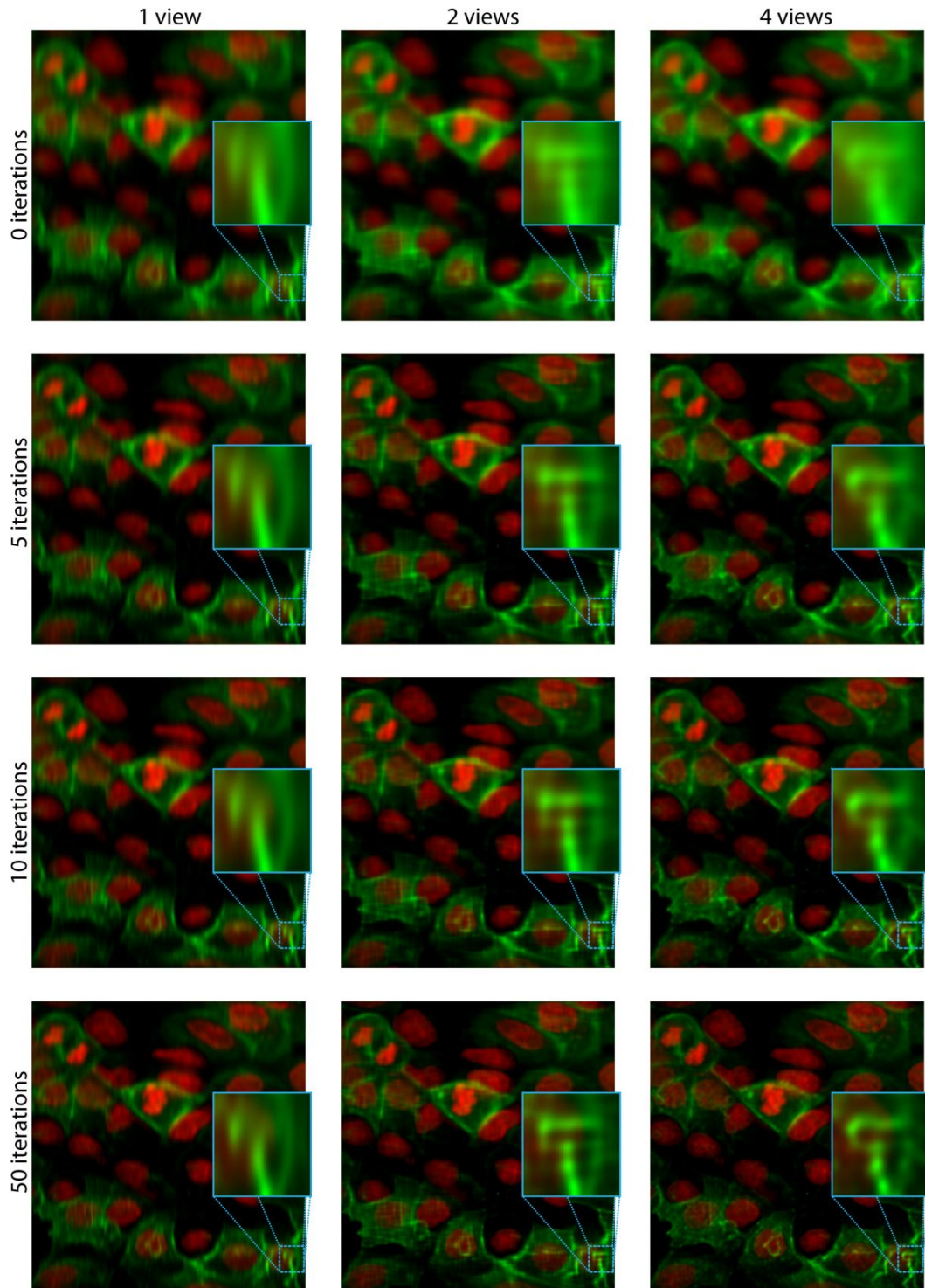


Figure 62: Fusion of simulated multiple-view images. The blurred images from Figure 61 were fused by the iterative expectation maximization algorithm. The images in this figure show how the number of different views (1,2 and 4) and number of algorithm iterations (0, 5, 10 and 50) effects the resulting image fusion. Please note the absence of horizontal patterns from single view images. However, the sharpness is generally improved by the deconvolution. Compare with the input images (Figure 61).

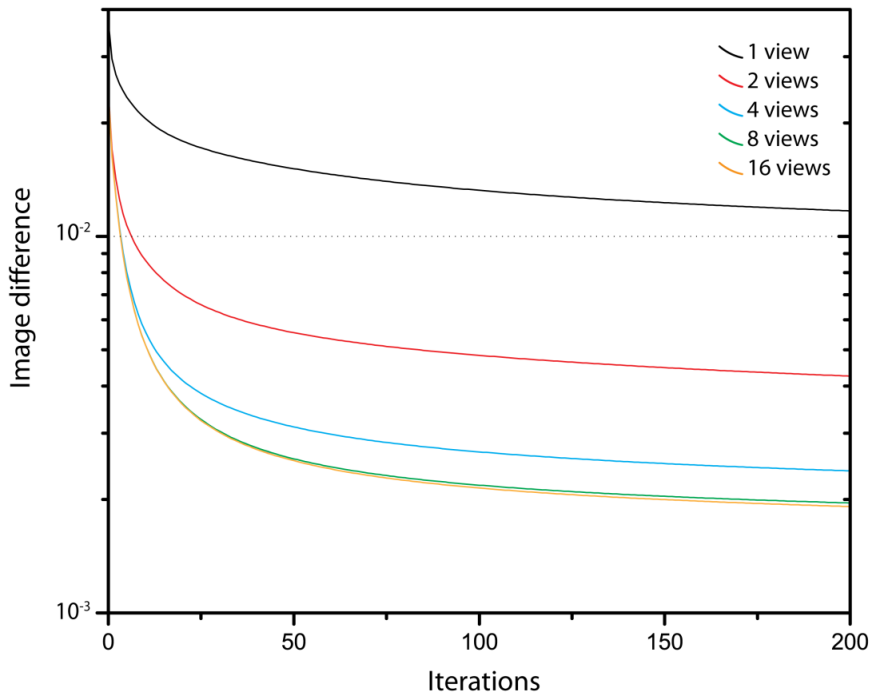


Figure 63: Image restoration by multiple-view fusion. Images, resulting from a fusion of a different number of simulated views (see extract in Figure 62) are compared with the original image (Figure 61, left). The difference between the fusion and the original image was measured by the normalized cross-correlation (NCC). The image difference indicates how close to unity (i.e. the perfect match) the NCC comes with increasing number of fusion iterations. The improvement is most rapid in the beginning of the fusion and when adding up to four views. The improvement provided by every view above eight views becomes negligible. Please note that this depends on the PSF elongation; in this simulation, PSF that is as long as it is wide, was used.

3.6 Examples of multiple-view microscopy on biological specimens

Iterative expectation-maximization algorithm was finally tested with LSFM images of a number of biological specimens. Figure 64 shows a high-magnification image of two budding-yeast cells. They expressed a GFP fused protein that associated with eisosomes, tiny cell surface structures. Fusion was produced from only two views (see also Figure 57).

Figure 65 and Figure 66 show multiple-view images of *Saccharomyces cerevisiae* yeast cells, labeled by Phalloidin-GFP, which associates with actin cables in the cell. Again, axial resolution in multiple-view images is significantly better than in the single views. Figure 66 shows also comparison of results of several different fusion algorithms described in section 3.5.4.

Multiple-view image of a medium sized specimen is shown in Figure 68. A large cellular spheroid of BXPC3 human pancreatic cancer cells labeled with DRAQ5 was imaged along 8 directions. Again, single views show only half of the whole object, but the fusion shows it all. Results of a single-view deconvolution and two different 8-view fusion algorithms are shown. Iterative expectation maximization fusion seems to provide the most convincing result in terms of sharpness and contrast. Figure 68 shows a detail from Figure 67. Note how well the iterative expectation maximization fusion reconstructs the image from relatively bad input images.

Figure 69 shows multiple-view image of a *Drosophila melanogaster* fruit fly head, as an example of a practically impenetrable specimen. Each of the single views only reveals roughly one quarter

of the specimen's surface. The multiple-view fusion, on the other hand, demonstrates a homogenous coverage of the entire surface. Similar results are obtained when a whole adult fruit fly is imaged (Figure 70).

Figure 71 shows an example of a large but fairly transparent specimen. Medaka fish (*Oryzias latipes*) was imaged along 8 directions. Single views exhibit continuous image degradation along the detection axis and reveal only approx. one half of the fish with reasonable contrast. The fused image, constructed from 8 views, covers the volume of the fish homogeneously, revealing the nervous system of the juvenile fish over its entire body.

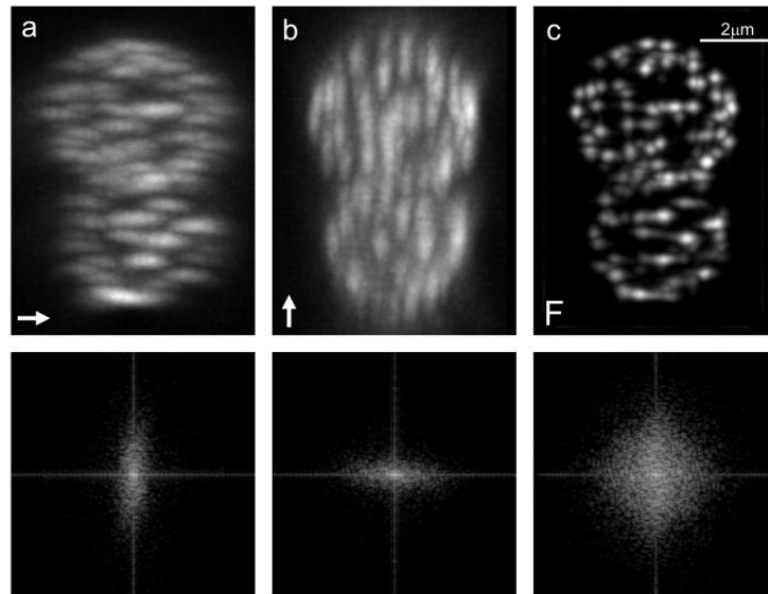


Figure 64: Multiple views of *Saccharomyces cerevisiae* expressing FIL1-GFP. The fluorophore distributions indicate the location of the protein FIL1, which is a part of the eisosomes, i.e. surface complexes involved in endocytosis. **a** and **b** show maximum intensity projections along the rotation axis through sets of images acquired along two orthogonal directions. The arrows indicate the directions of the detection axes. **c** shows the fusion of the two images shown in **a** and **b**. The bottom row shows the corresponding power spectra. The original data set has a size of 200x200x50 volume elements. The image stacks were recorded with a Zeiss Achroplan 100x/1.0W, the excitation wavelength was 0.488µm and the emission was recorded above 0.510µm. The scale bar in **c** is also valid for **a** and **b**. The sample was provided by Marko Kaksonen (EMBL).

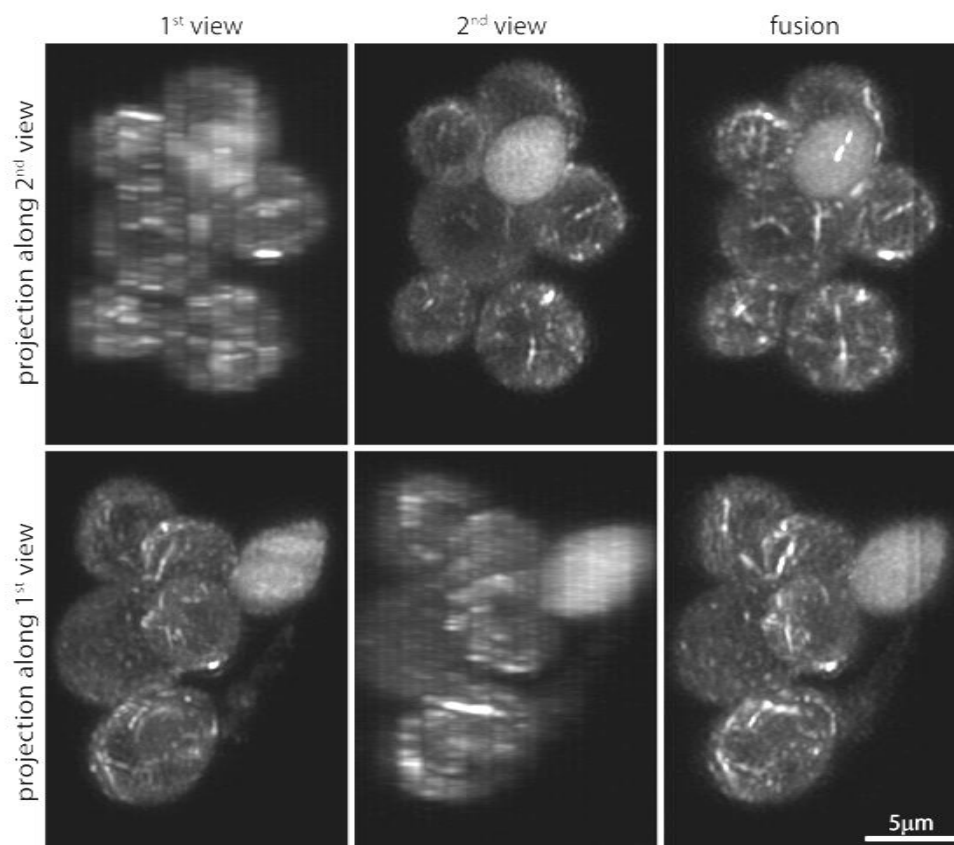


Figure 65: Multiple views of *Saccharomyces cerevisiae* labeled with Phalloidin-GFP (marks actin cables).

The first two columns show two individual multiple-view images acquired along two orthogonal directions. The column on the right shows the results of the fusion of 12 such stacks, using the iterative expectation maximization algorithm. The first row of images shows the maximum-intensity projections along a direction, which is parallel to the detection axis of the view in the left column, and the second row an orthogonal projection, which is parallel to the detection axis of the view in the second column. The lateral projections of the individual views provide a good level of detail, while the axial views suffer from poor resolution and sparse sampling (only 9 planes are acquired along the whole lump of cells). The fusion, on the other hand, provides a good level of detail regardless of the projection direction. Please note how the fusion images fit the lateral projections of the individual views. The images were recorded with a SPIM and Zeiss Achroplan 100x/1.0W objective lens, the excitation wavelength was $0.488\mu\text{m}$ and the emission was recorded above $0.510\mu\text{m}$. Specimen was provided by C. Taxis (University of Marburg) and M. Knop (EMBL Heidelberg).

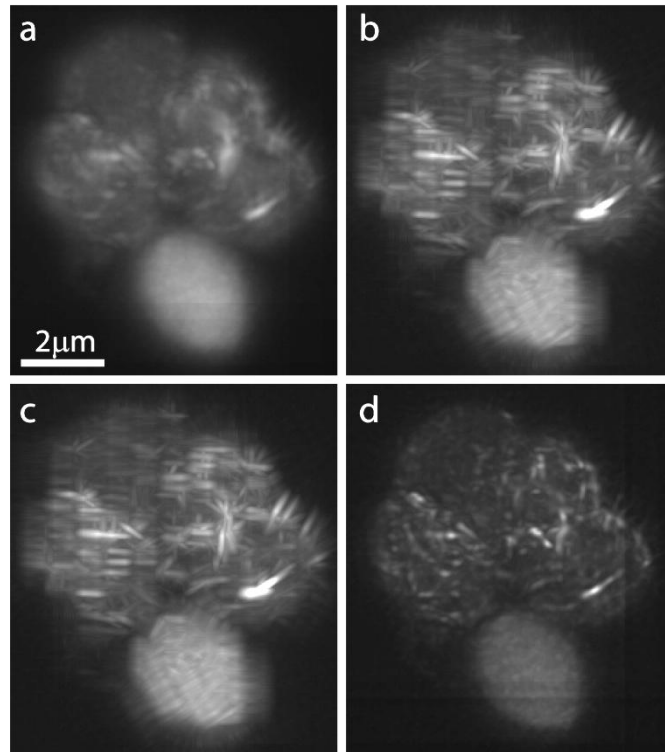


Figure 66: Comparison of different fusion techniques – transparent specimen. The multiple-view images from Figure 65 were fused using different fusion algorithms: **a** mean of all the views (section 3.5.4.10), **b** maximum intensity (section 3.5.4.2), **c** average, weighted by blurred first derivative (section 3.5.4.2), **d** iterative expectation maximization algorithm (section 3.5.4.14). Only the last algorithm provides a clear gain in sharpness. The algorithms used for **b** and **c** clearly preserve the elongated shapes produced by the anisotropic PSFs of the input stacks. Image in **a** is more isotropic, but considerably less sharp. See Figure 65 for imaging details. Specimen was provided by C. Taxis (University of Marburg) and M. Knop (EMBL Heidelberg).

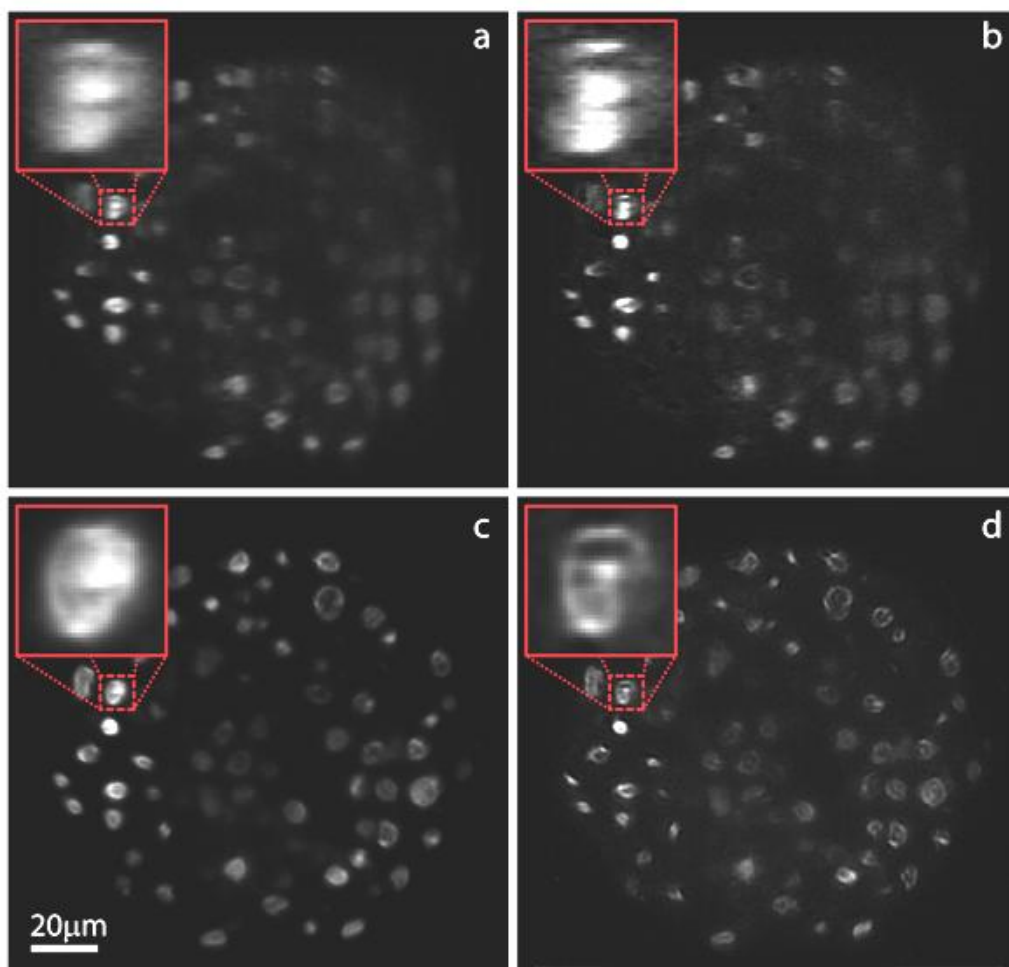


Figure 67: Comparison of different fusion techniques – semi-transparent specimen. The input images from Figure 51 were fused using a first-derivative weighted average and by the iterative expectation maximization algorithm. **a** One of the eight input views, used for the fusion. **b** The result of a deconvolution; the sharpness is improved but the definition along the vertical direction remains poor. **c** The images were fused using a weighted average. The weights were proportional to local image intensity derivative, blurred with a 3×3 box kernel (section 99). **d** The result of the iterative expectation maximization fusion. See Figure 51 for imaging details.

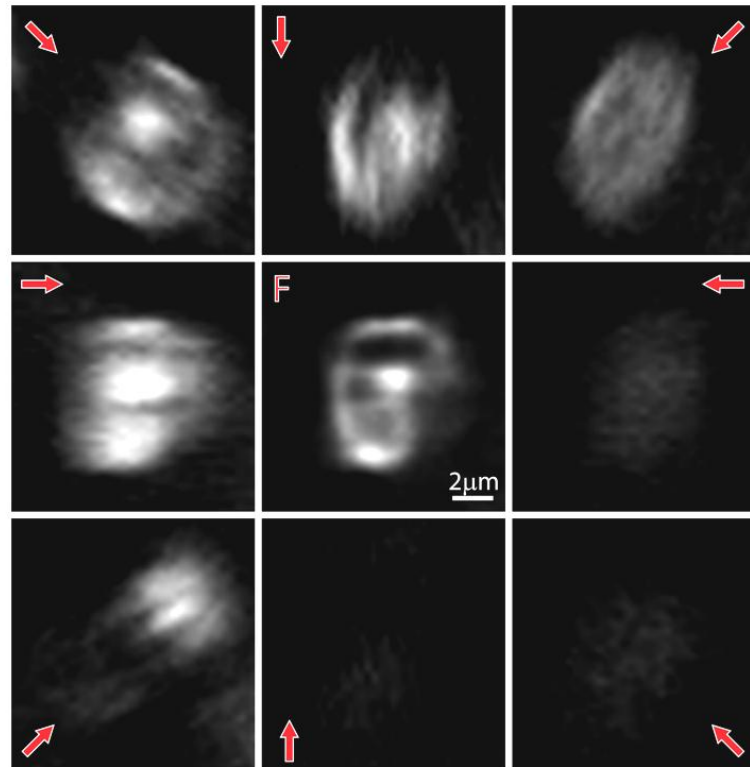


Figure 68: Input images and their fusion. The images show the highlighted nucleus from Figure 51 (left nucleus) and Figure 67. The center image is a fusion of the eight views that are distributed around it, according to the relative orientation of their detection axes (indicated by the arrows). The quality of the nucleus' image is influenced by the amount of tissue the light has to pass to and fro the nucleus (see Figure 67). The three images in the bottom-right corner are almost entirely indiscernible from the background. The bottom-left and top-right images are badly degraded as well. Nevertheless, the fusion creates a vastly improved image of the nucleus from an apparently poor pool of data. See Figure 51 for imaging details.

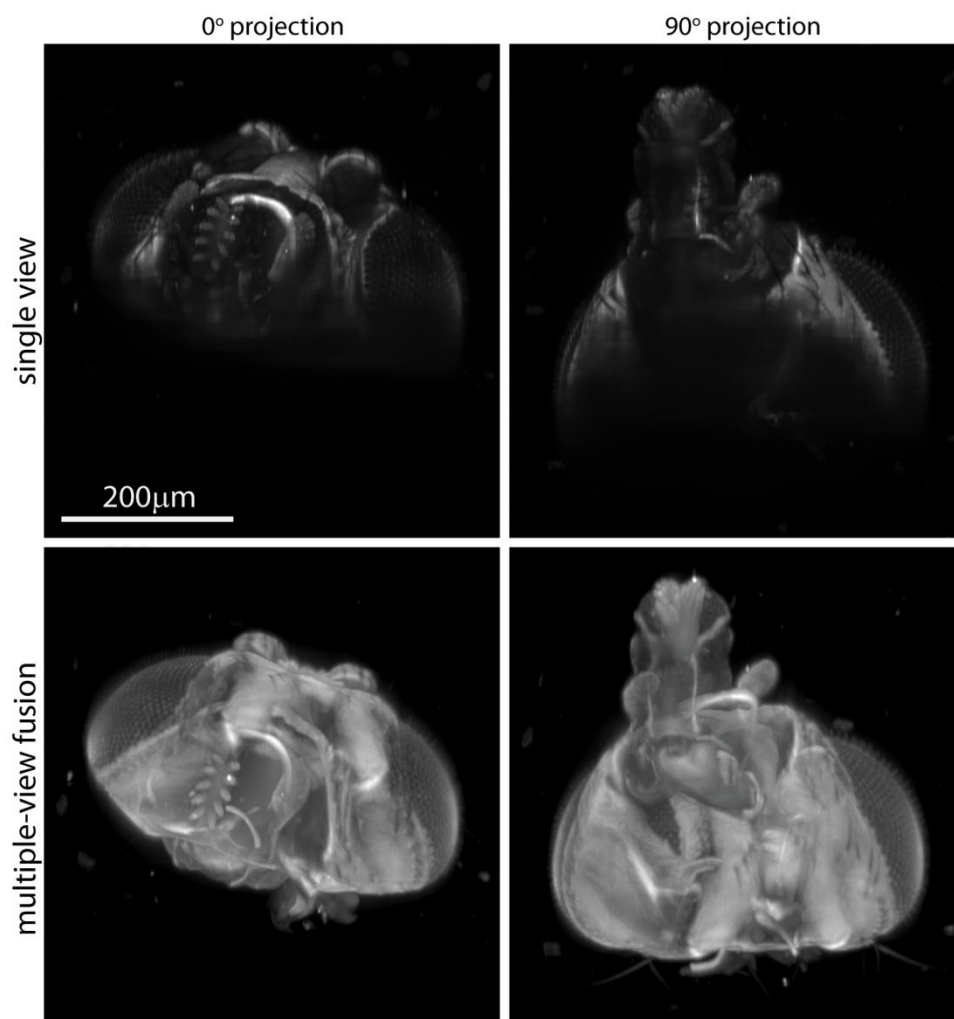


Figure 69: Multiple-view image of a fully opaque specimen. The images show maximum parallel projections of three-dimensional *Drosophila melanogaster* fruit fly head images. The top row shows two single views acquired along two orthogonal directions (detection from above). The bottom row shows result of a multiple-view fusion of six such views. Images in the second row have the same orientation as the corresponding images in the first row. While each of the single views only reveals approx. one quarter of the object's surface, the fused image covers the surface homogeneously. Autofluorescence of a non-treated wild-type fly was recorded by SPIM and Carl Zeiss Fluor 5x/0.25 detection lens. Illumination wavelength was 488nm, emission above 488nm was detected.

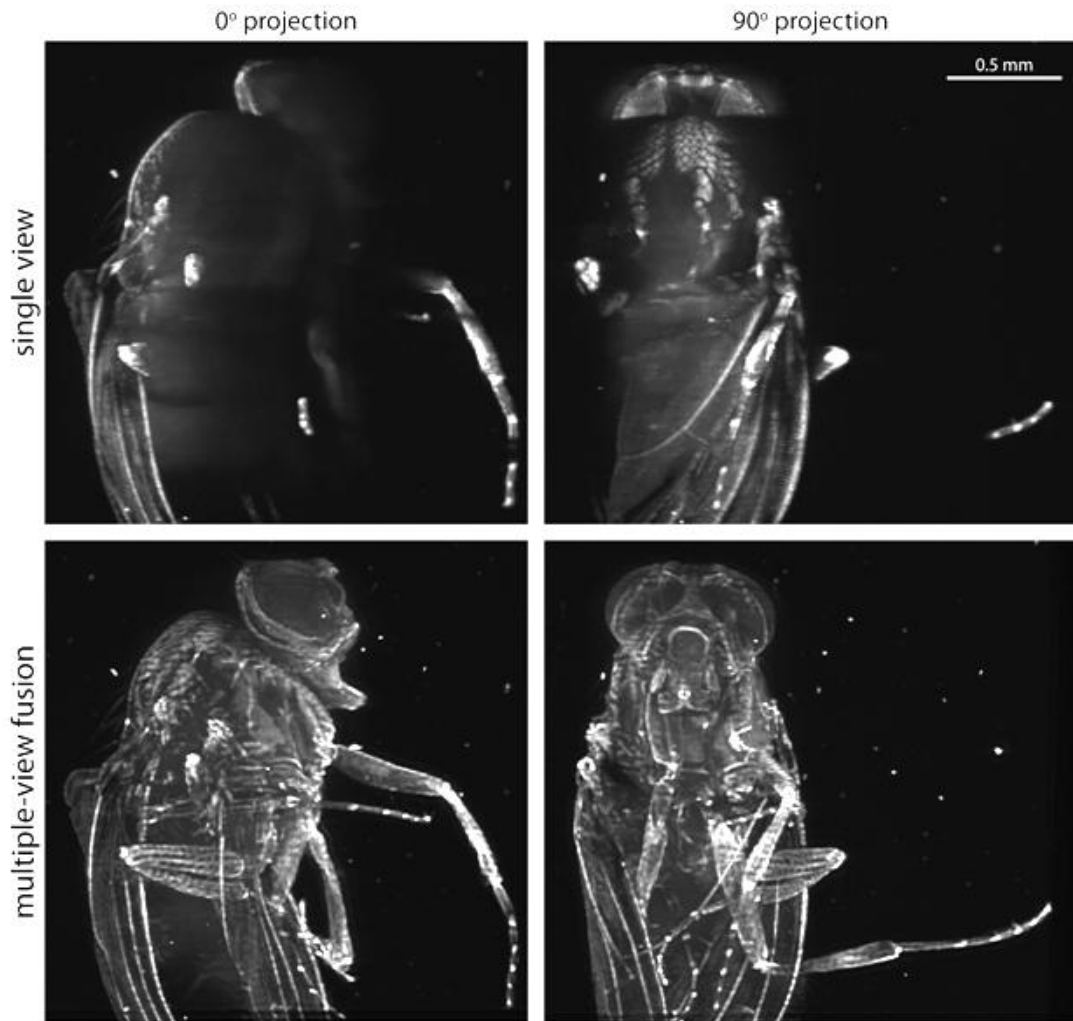


Figure 70: Multiple-view image of a large and opaque specimen's surface. The images show maximum parallel projections of three-dimensional *Drosophila melanogaster* adult fly images. The top row shows a single view projected along two orthogonal directions. The bottom row shows result of a multiple-view fusion of twelve views. Images in the second row have the same orientation as the corresponding images in the first row. Autofluorescence of a non-treated wild-type fly was recorded by SPIM and Carl Zeiss Fluar 2.5x/0.12 detection lens. Illumination wavelength was 488nm, while emission above 488nm was detected.

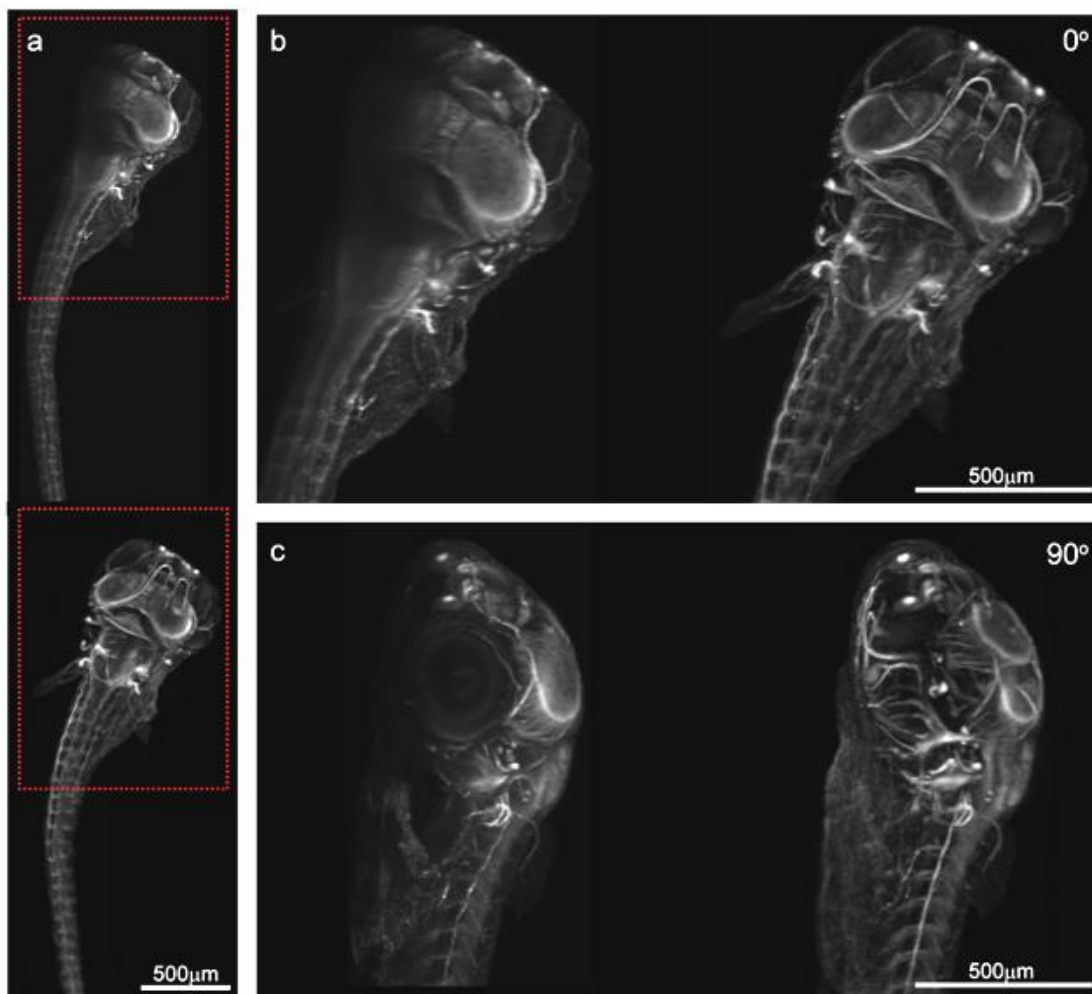


Figure 71: Single view and multiple-view fusion of juvenile Medaka fish with an acetylated tubulin immuno-staining. **a** shows a maximum intensity projection of a single view (top) and a fusion of 8 views (bottom). **b** and **c** show the magnified area indicated by the rectangle in **a**, projected along two orthogonal directions. The images on the left in **b** and **c** show a single view, while the images on the right show a fusion of 8 views. Each of the original data sets has a size of 500x500x1000 picture elements. The image stacks were recorded with a Zeiss Fluar 2.5x/0.12 objective, the excitation wavelength was 0.488µm and the emission was recorded above 0.488µm. Preparation by Annette Schmidt and Lazaro Centanin, imaging by Philipp Keller (EMBL).

4 SUMMARY AND OUTLOOK

Light microscopy has been the most popular imaging method used by the life sciences for over a century. Its main advantages over other imaging methods (such as electron microscopy, x-ray tomography and NMR imaging) are: relatively cheap instrumentation, simple specimen preparation, the configurable signal specificity offered by fluorescent labeling and the possibility of imaging living specimens under physiologically relevant conditions. This thesis presents the light-sheet based fluorescence microscope (LSFM), a new microscopy method that is expected to have a significant impact on modern biological imaging. LSFM exposes a fluorescent specimen to much smaller doses of energies than other common fluorescence microscopes, e.g. up to approx. 200 times less energy than an epifluorescence microscope and up to 5000 times less energy than a confocal laser scanning microscope. Furthermore, a specimen in a LSFM is mounted in a three-dimensional environment that more resembles the physiological environment than the flat glass substrata used by other common microscopes. Finally, due to the LSFM's illumination based optical sectioning, LSFM's image contrast is far superior to that of other common fluorescence microscopes, images demonstrate an improved dynamic range and a resolution comparable to that of confocal microscopes. Based on these impressive features, the potential that the LSFM has to have a dramatic impact on developmental and cell biology is clear.

In future, LSFM's imaging capabilities will probably be combined with various methods for optical specimen manipulation and expanded to different imaging modalities. A few examples of the countless possibilities have already been implemented at EMBL's light microscopy group in Heidelberg: UV based laser cutter, structured illumination and FLIM (fluorescence lifetime imaging microscopy). Another prospect, enabled by LSFM's high contrast, is high-precision fluorescence localization microscopy (e.g. light-sheet based PALM or STORM).

LSFM also seems better suited for multiple-view imaging than any other fluorescence microscope. In this thesis, the foundations of LSFM based multiple-view microscopy are laid. Acquisition of a multiple-view set is discussed and demonstrated by a number of specimens imaged using the EMBL's LSFM implementations. Image processing methods for registration and fusion of multiple views are evaluated. Previously reported ideas for fusion of multiple views are reviewed and original solutions are proposed and implemented. The potential and limits of multiple-view microscopy with LSFM is tested using a number of biological specimens, ranging in size from single yeast (*Saccharomyces cerevisiae*) cells to cellular spheroids, embryos (*Drosophila melanogaster*, *Danio rerio*) and adult insects.

Multiple-view microscopy has the potential to improve biological imaging over the entire range of specimen sizes. Scientists dealing with small specimens can benefit primarily from the improved resolution offered by multiple views. However, multiple-view microscopy also has the capability of creating complete images of opaque specimens that are too large to be imaged adequately well along any single direction. This ability is of utmost importance for developmental biology. Penetration depth provided by most modern fluorescence microscopes is in many cases insufficient to reveal complex cell movements involved in embryonic and tissue development. Multiple-view microscopy might prove to be the optimal solution to this problem.

The full potentials of the multiple-view microscopy will only be unlocked if the time required for multiple-view image processing can be drastically reduced. Ideally, images should be fused in real time, i.e. as they are recorded. This thesis discusses algorithms that are limited to relatively simple image processing methods, they are implemented in real space (as opposed to Fourier or frequency domain) and can thus be easily parallelized. Unfortunately, a modern personal computer currently requires several hours to process a typical multiple-view dataset. This is likely to change in the near future, considering the rapid development of faster and multi core processors, which could make the multiple-view microscopy a practical and popular method for biological imaging.

BIBLIOGRAPHY

- [1] F. Pampaloni, E.G. Reynaud, and E.H.K. Stelzer, "The third dimension bridges the gap between cell culture and live tissue," *Nature Reviews Molecular Cell Biology*, vol. 8, Oct. 2007, pp. 839-845.
- [2] E.G. Reynaud, U. Kržič, K. Greger, and E.H.K. Stelzer, "Light sheet-based fluorescence microscopy: more dimensions, more photons, and less photodamage," *HFSP Journal*, vol. 2, Oct. 2008, pp. 266-275.
- [3] A. Rohrbach, C. Tischer, D. Neumayer, E. Florin, and E.H.K. Stelzer, "Trapping and tracking a local probe with a photonic force microscope," *Review of Scientific Instruments*, vol. 75, Jun. 2004, pp. 2197-2210.
- [4] J. Colombelli, E.G. Reynaud, and E.H.K. Stelzer, "Investigating relaxation processes in cells and developing organisms: from cell ablation to cytoskeleton nanosurgery," *Methods in Cell Biology*, vol. 82, 2007, pp. 267-91.
- [5] J.W. Lichtman and J. Conchello, "Fluorescence microscopy," *Nature Methods*, vol. 2, Dec. 2005, pp. 910-9.
- [6] N.C. Shaner, P.A. Steinbach, and R.Y. Tsien, "A guide to choosing fluorescent proteins," *Nature methods*, vol. 2, Dec. 2005, pp. 905-9.
- [7] B.N.G. Giepmans, S.R. Adams, M.H. Ellisman, and R.Y. Tsien, "The Fluorescent Toolbox for Assessing Protein Location and Function," *Science*, vol. 312, Apr. 2006, pp. 217-224.
- [8] Invitrogen, Molecular Probes, (undated) [online] "The Handbook - A Guide to Fluorescent Probes and Labeling Techniques," Viewed Apr. 2008, Available: <http://www.invitrogen.com/site/us/en/home/References/Molecular-Probes-The-Handbook.html>.
- [9] A.H. Coons, H.J. Creech, and R.N. Jones, "Immunological properties of an antibody containing a fluorescent group," *Proceedings of the Society for Experimental Biology and Medicine*, vol. 47, 1941, pp. 200-202.
- [10] A. Miyawaki, O. Griesbeck, R. Heim, and R.Y. Tsien, "Dynamic and quantitative Ca²⁺ measurements using improved cameleons," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 96, Mar. 1999, pp. 2135-40.
- [11] M. Terasaki, L. Loew, J. Lippincott-Schwartz, and K. Zaal, "Fluorescent staining of subcellular organelles: ER, Golgi complex, and mitochondria," *Current Protocols in Cell Biology*, vol. 4.4, May. 2001.
- [12] J. Lippincott-Schwartz and G.H. Patterson, "Development and Use of Fluorescent Protein Markers in Living Cells," *Science*, vol. 300, Apr. 2003, pp. 87-91.
- [13] M. Chalfie, Y. Tu, G. Euskirchen, W. Ward, and D. Prasher, "Green fluorescent protein as a marker for gene expression," *Science*, vol. 263, Feb. 1994, pp. 802-805.
- [14] Y.A. Labas, N.G. Gurskaya, Y.G. Yanushevich, A.F. Fradkov, K.A. Lukyanov, S.A. Lukyanov, and M.V. Matz, "Diversity and evolution of the green fluorescent protein

- family,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, Apr. 2002, pp. 4256-4261.
- [15] H. Morise, O. Shimomura, F.H. Johnson, and J. Winant, “Intermolecular energy transfer in the bioluminescent system of *Aequorea*,” *Biochemistry*, vol. 13, Jun. 1974, pp. 2656-62.
- [16] N.C. Shaner, R.E. Campbell, P.A. Steinbach, B.N.G. Giepmans, A.E. Palmer, and R.Y. Tsien, “Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein,” *Nature Biotechnology*, vol. 22, 2004, pp. 1567-1572.
- [17] R.F. Service, “Nobel Prize in Chemistry: Three Scientists Bask in Prize's Fluorescent Glow,” *Science*, vol. 322, Oct. 2008, p. 361.
- [18] K.A. Lukyanov, D.M. Chudakov, S. Lukyanov, and V.V. Verkhusha, “Innovation: Photoactivatable fluorescent proteins,” *Nature Reviews Molecular Cell Biology*, vol. 6, Nov. 2005, pp. 885-91.
- [19] G.H. Patterson and J. Lippincott-Schwartz, “A Photoactivatable GFP for Selective Photolabeling of Proteins and Cells,” *Science*, vol. 297, Sep. 2002, pp. 1873-1877.
- [20] R. Ando, H. Hama, M. Yamamoto-Hino, H. Mizuno, and A. Miyawaki, “An optical marker based on the UV-induced green-to-red photoconversion of a fluorescent protein,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, Oct. 2002, pp. 12651-12656.
- [21] S. Habuchi, R. Ando, P. Dedecker, W. Verheijen, H. Mizuno, A. Miyawaki, and J. Hofkens, “Reversible single-molecule photoswitching in the GFP-like fluorescent protein Dronpa,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, Jul. 2005, pp. 9511-6.
- [22] J. Wiedenmann, S. Ivanchenko, F. Oswald, F. Schmitt, C. Rucker, A. Salih, K. Spindler, and G.U. Nienhaus, “EosFP, a fluorescent marker protein with UV-inducible green-to-red fluorescence conversion,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, Nov. 2004, pp. 15905-15910.
- [23] R. Ando, H. Mizuno, and A. Miyawaki, “Regulated Fast Nucleocytoplasmic Shuttling Observed by Reversible Protein Highlighting,” *Science*, vol. 306, Nov. 2004, pp. 1370-1373.
- [24] C. Seydel, “Quantum Dots Get Wet,” *Science*, vol. 300, Apr. 2003, pp. 80-81.
- [25] A.P. Alivisatos, W. Gu, and C. Larabell, “Quantum dots as cellular probes,” *Annual Review of Biomedical Engineering*, vol. 7, 2005, pp. 55-76.
- [26] B. Ballou, B.C. Lagerholm, L.A. Ernst, M.P. Bruchez, and A.S. Waggoner, “Noninvasive imaging of quantum dots in mice,” *Bioconjugate Chemistry*, vol. 15, pp. 79-86.
- [27] I.J. Cox, “Scanning optical fluorescence microscopy,” *Journal of Microscopy*, vol. 133, Feb. 1984, pp. 149-54.
- [28] S. Grill and E.H.K. Stelzer, “Method to calculate lateral and axial gain factors of optical setups with a large solid angle,” *Journal of the Optical Society of America A*, vol. 16, Nov. 1999, pp. 2658-2665.
- [29] E.H.K. Stelzer and S. Grill, “The uncertainty principle applied to estimate focal spot dimensions,” *Optics Communications*, vol. 173, Jan. 2000, pp. 51-56.
- [30] M. Born and E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*, Pergamon Press, 1980.
- [31] E.H.K. Stelzer, “Contrast, resolution, pixelation, dynamic range and signal-to-noise ratio: fundamental limits to resolution in fluorescence light microscopy,” *Journal of Microscopy*, vol. 189, 1998, pp. 15-24.

- [32] J.E.N. Jonkman and E.H.K. Stelzer, "Resolution and contrast in confocal and two-photon microscopy," *Confocal and Two-Photon Microscopy: Foundations, Applications, and Advances*, A. Diaspro, ed., New York: Wiley-Liss, 2002, pp. 101–125.
- [33] J. Conchello and J.W. Lichtman, "Optical sectioning microscopy," *Nature Methods*, vol. 2, Dec. 2005, pp. 920-931.
- [34] S.W. Hell, K. Bahlmann, M. Schrader, A. Soini, H. Malak, I. Gryczynski, and J.R. Lakowicz, "Three-Photon Excitation in Fluorescence Microscopy," *Journal of Biomedical Optics*, vol. 1, 1996, p. 71.
- [35] S.W. Hell, J. Soukka, and P.E. Hänninen, "Two- and multiphoton detection as an imaging mode and means of increasing the resolution in far-field light microscopy," *Bioimaging*, vol. 3, 1995, pp. 65-69.
- [36] M. Göppert-Mayer, "Über Elementarakte mit zwei Quantensprüngen," *Annalen der Physik*, vol. 401, 1931, pp. 273-294.
- [37] F. Helmchen and W. Denk, "Deep tissue two-photon microscopy," *Nature Methods*, vol. 2, Dec. 2005, pp. 932-40.
- [38] J. Mertz, "Nonlinear microscopy: new techniques and applications," *Current Opinion in Neurobiology*, vol. 14, Oct. 2004, pp. 610-6.
- [39] P.J. Campagnola and L.M. Loew, "Second-harmonic imaging microscopy for visualizing biomolecular arrays in cells, tissues and organisms," *Nature Biotechnology*, vol. 21, Nov. 2003, pp. 1356-60.
- [40] B.E. Saleh and M.C. Teich, *Fundamentals of Photonics*, Wiley & Sons, 2007.
- [41] H. Wang, Y. Fu, P. Zickmund, R. Shi, and J. Cheng, "Coherent Anti-Stokes Raman Scattering Imaging of Axonal Myelin in Live Spinal Tissues," *Biophysical Journal*, vol. 89, Jul. 2005, pp. 581-591.
- [42] J. Cheng, Y.K. Jia, G. Zheng, and X.S. Xie, "Laser-Scanning Coherent Anti-Stokes Raman Scattering Microscopy and Applications to Cell Biology," *Biophysical Journal*, vol. 83, Jul. 2002, pp. 502-509.
- [43] J.X. Cheng, A. Volkmer, and X.S. Xie, "Theoretical and experimental characterization of coherent anti-Stokes Raman scattering microscopy," *Journal of the Optical Society of America B*, vol. 19, 2002, pp. 1363-1375.
- [44] D. Stephens, *Cell Imaging: Methods Express Series*, Scion Publishing Ltd., 2005.
- [45] M.A.A. Neil, R. Juskaitis, and T. Wilson, "Method of obtaining optical sectioning by using structured light in a conventional microscope," *Optics Letters*, vol. 22, Dec. 1997, pp. 1905-1907.
- [46] E.H.K. Stelzer and S. Lindek, "Fundamental reduction of the observation volume in far-field light microscopy by detection orthogonal to the illumination axis: confocal theta microscopy," *Optics Communications*, vol. 111, Oct. 1994, pp. 536-547.
- [47] F. Haar, J. Swoger, and E.H. Stelzer, "Developments and applications of confocal theta microscopy," *Proceedings of SPIE*, vol. 3605, May. 1999, pp. 48-54.
- [48] J. Swoger, S. Lindek, T. Stefany, F. Haar, and E.H.K. Stelzer, "A confocal fiber-coupled single-lens theta microscope," *Review of Scientific Instruments*, vol. 69, 1998, pp. 2956-2963.
- [49] S. Lindek, T. Stefany, and E.H.K. Stelzer, "Single-lens theta microscopy—a new implementation of confocal theta microscopy," *Journal of Microscopy*, vol. 188, 1997, pp. 280-284.
- [50] B. Bailey, D.L. Farkas, D.L. Taylor, and F. Lanni, "Enhancement of axial resolution in fluorescence microscopy by standing-wave excitation," *Nature*, vol. 366, Nov. 1993, pp. 44-48.

- [51] M.G. Gustafsson, D.A. Agard, and J.W. Sedat, "I5M: 3D widefield light microscopy with better than 100 nm axial resolution," *Journal of microscopy*, vol. 195, Jul. 1999, pp. 10-6.
- [52] S. Hell and E.H.K. Stelzer, "Properties of a 4Pi confocal fluorescence microscope," *Journal of the Optical Society of America A*, vol. 9, Dec. 1992, p. 2159.
- [53] S. Hell and E.H.K. Stelzer, "Fundamental improvement of resolution with a 4Pi-confocal fluorescence microscope using two-photon excitation," *Optics Communications*, vol. 93, Oct. 1992, pp. 277-282.
- [54] S. Lindek, R. Pick, and E.H.K. Stelzer, "Confocal theta microscope with three objective lenses," *Review of Scientific Instruments*, vol. 65, Nov. 1994, pp. 3367-3372.
- [55] S. W. Hell, "Double-confocal microscope," U.S. Patent 0491289, 1990.
- [56] S.W. Hell, "Toward fluorescence nanoscopy," *Nature Biotechnology*, vol. 21, 2003, pp. 1347-1355.
- [57] E. Abbe, "Beitrage zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung," *Archiv für mikroskopische Anatomie*, vol. 9, 1873, pp. 413-420.
- [58] E.H. Synge, "A suggested method for extending microscopic resolution into the ultramicroscopic region," *Philosophical Magazine*, vol. 6, 1928, p. 356.
- [59] E.A. Ash and G. Nicholls, "Super-resolution aperture scanning microscope," *Nature*, vol. 237, Jun. 1972, pp. 510-2.
- [60] D.W. Pohl, W. Denk, and M. Lanz, "Optical stethoscopy: Image recording with resolution $\lambda/20$," *Applied Physics Letters*, vol. 44, Apr. 1984, pp. 651-653.
- [61] A. Lewis, M. Isaacson, A. Harootunian, and A. Muray, "Development of a 500 Å spatial resolution light microscope," *Ultramicroscopy*, vol. 13, 1984, pp. 227-231.
- [62] D. Courjon, *Near Field Microscopy and Near Field Optics*, Imperial College Press, 2003.
- [63] E. Betzig and R.J. Chichester, "Single Molecules Observed by Near-Field Scanning Optical Microscopy," *Science*, vol. 262, Nov. 1993, pp. 1422-1425.
- [64] M.G. Gustafsson, "Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy," *Journal of microscopy*, vol. 198, May. 2000, pp. 82-7.
- [65] E. Betzig, G.H. Patterson, R. Sougrat, O.W. Lindwasser, S. Olenych, J.S. Bonifacino, M.W. Davidson, J. Lippincott-Schwartz, and H.F. Hess, "Imaging Intracellular Fluorescent Proteins at Nanometer Resolution," *Science*, Aug. 2006, pp. 1642-45.
- [66] R.E. Thompson, D.R. Larson, and W.W. Webb, "Precise Nanometer Localization Analysis for Individual Fluorescent Probes," *Biophysical Journal*, vol. 82, May. 2002, pp. 2775-2783.
- [67] B. Huang, W. Wang, M. Bates, and X. Zhuang, "Three-Dimensional Super-Resolution Imaging by Stochastic Optical Reconstruction Microscopy," *Science*, vol. 319, Feb. 2008, pp. 810-813.
- [68] S.W. Hell and J. Wichmann, "Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy," *Optics Letters*, vol. 19, 1994, pp. 780-780.
- [69] T.A. Klar, S. Jakobs, M. Dyba, A. Egner, and S.W. Hell, "Fluorescence microscopy with diffraction resolution barrier broken by stimulated emission," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 97, Jul. 2000.
- [70] G. Donnert, J. Keller, R. Medda, M.A. Andrei, S.O. Rizzoli, R. Lührmann, R. Jahn, C. Eggeling, and S.W. Hell, "Macromolecular-scale resolution in biological fluorescence microscopy," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, Aug. 2006, pp. 11440-5.

- [71] Sebastian Enders, "Entwicklung und Erprobung eines Theta-Großfeld-Mikroskops," Diploma thesis, Ruprecht-Karls-Universität Heidelberg, 2003.
- [72] Henry Siedentopf and Richard Adolf Zsigmondy, "Ueber Sichtbarmachung ultramikroskopischer Teilchen, mit besonderer Anwendung auf Goldrubingläser," *Drude's Annalen der Physik*, vol. 315, 1903, pp. 1-39.
- [73] R. Zsigmondy, *Colloids And The Ultramicroscope*, Read Books, 2007.
- [74] E. Fuchs, J. Jaffe, R. Long, and F. Azam, "Thin laser light sheet microscope for microbial oceanography," *Optics Express*, vol. 10, Jan. 2002, pp. 145-154.
- [75] Peter Zampol, "Method of photography," U.S. Patent 2928734, March 1960.
- [76] Dan McLachlan, "Microscope," U.S. Patent 3398634, August 1968.
- [77] W. Sharp and C. Kazilek, "Scanning Light Photomacrography," *Darkroom and Creative Camera Techniques*, 1990, pp. 43-45.
- [78] D. Huber, M. Keller, and D. Robert, "3D light scanning macrography," *Journal of Microscopy*, vol. 203, Aug. 2001, pp. 208-13.
- [79] A.H. Voie, D.H. Burns, and F.A. Spelman, "Orthogonal-plane fluorescence optical sectioning: three-dimensional imaging of macroscopic biological specimens," *Journal of Microscopy*, vol. 170, Jun. 1993, pp. 229-36.
- [80] J. Huisken, J. Swoger, F. Del Bene, J. Wittbrodt, and E.H.K. Stelzer, "Optical Sectioning Deep Inside Live Embryos by Selective Plane Illumination Microscopy," *Science*, vol. 305, Aug. 2004, pp. 1007-1009.
- [81] K. Greger, J. Swoger, and E.H.K. Stelzer, "Basic building units and properties of a fluorescence single plane illumination microscope," *Review of Scientific Instruments*, vol. 78, Feb. 2007, p. 23705.
- [82] C.J. Engelbrecht, K. Greger, E.G. Reynaud, U. Kržič, J. Colombelli, and E.H. Stelzer, "Three-dimensional laser microsurgery in light-sheet based microscopy (SPIM)," *Optics Express*, vol. 15, May. 2007, pp. 6420-6430.
- [83] T. Breuninger, K. Greger, and E.H.K. Stelzer, "Lateral modulation boosts image quality in single plane illumination fluorescence microscopy," *Optics Letters*, vol. 32, Jul. 2007, pp. 1938-40.
- [84] M. Neetz, "Light sheet based frequency domain fluorescence lifetime imaging," Diploma thesis, Ruprecht-Karls-Universität Heidelberg, 2008.
- [85] CVI Melles Griot, (undated), [online], "Optics Guide," Viewed Feb 2007, Available: <http://www.mellesgriot.com/products/optics/toc.htm>
- [86] S. Inoue, *Video Microscopy*, Plenum Publishing Corporation, 1986.
- [87] Tobias Breuninger, "Optimierung eines Lichtscheibenmikroskops durch laterale Modulation zur Verbesserung der Bildqualität und Charakterisierung einer 3D-Zellkultur aus MDCK-Zellen, kultiviert in verschiedenen Gel-Matrices," Master thesis, Hochschule Albstadt-Sigmaringen, 2006.
- [88] P.J. Keller, A.D. Schmidt, J. Wittbrodt, and E.H.K. Stelzer, "Reconstruction of Zebrafish Early Embryonic Development by Scanned Light Sheet Microscopy," *Science*, Nov. 2008, pp. 1065-9.
- [89] F. Pampaloni, E.G. Reynaud, and E.H.K. Stelzer, "The third dimension bridges the gap between cell culture and live tissue," *Nature Reviews Molecular Cell Biology*, vol. 8, Oct. 2007, pp. 839-845.
- [90] J. Swoger, P. Verveer, K. Greger, J. Huisken, and E.H.K. Stelzer, "Multi-view image fusion improves resolution in three-dimensional microscopy," *Optics Express*, vol. 15, Jun. 2007, pp. 8029-8042.

- [91] C. Taxis, C. Maeder, S. Reber, N. Rathfelder, K. Miura, K. Greger, E.H.K. Stelzer, and M. Knop, "Dynamic Organization of the Actin Cytoskeleton During Meiosis and Spore Formation in Budding Yeast," *Traffic*, vol. 7, 2006, pp. 1628-1642.
- [92] P.J. Keller, F. Pampaloni, G. Lattanzi, and E.H.K. Stelzer, "Three-Dimensional Microtubule Behavior in Xenopus Egg Extracts Reveals Four Dynamic States and State-Dependent Elastic Properties," *Biophysical Journal*, vol. 95, Aug. 2008, pp. 1474-1486.
- [93] P.J.M. Van Haastert and P.N. Devreotes, "Chemotaxis: signalling the way forward," *Nature Reviews Molecular Cell Biology*, vol. 5, Aug. 2004, pp. 626-34.
- [94] M. Vicente-Manzanares, D.J. Webb, and A.R. Horwitz, "Cell migration at a glance," *Journal of Cell Science*, vol. 118, Nov. 2005, pp. 4917-9.
- [95] D. Chodniewicz and R.L. Klemke, "Guiding cell migration through directed extension and stabilization of pseudopodia," *Experimental Cell Research*, vol. 301, Nov. 2004, pp. 31-7.
- [96] P. Haas and D. Gilmour, "Chemokine signaling mediates self-organizing tissue migration in the zebrafish lateral line," *Developmental Cell*, vol. 10, May. 2006, pp. 673-80.
- [97] C. Ribeiro, V. Petit, and M. Affolter, "Signaling systems, guided cell migration, and organogenesis: insights from genetic studies in Drosophila," *Developmental Biology*, vol. 260, Aug. 2003, pp. 1-8.
- [98] R.J. Greenspan, *Fly Pushing: The Theory and Practice of Drosophila Genetics*, Cold Spring Harbor Laboratory Press, 1997.
- [99] A. Jacinto, S. Woolner, and P. Martin, "Dynamic analysis of dorsal closure in Drosophila: from genetics to cell biology," *Developmental Cell*, vol. 3, Jul. 2002, pp. 9-19.
- [100] D.J. Montell, "Border-cell migration: the race is on," *Nature Reviews Molecular Cell Biology*, vol. 4, Jan. 2003, pp. 13-24.
- [101] A. Uv, R. Cantera, and C. Samakovlis, "Drosophila tracheal morphogenesis: intricate cellular solutions to basic plumbing problems," *Trends in Cell Biology*, vol. 13, Jun. 2003, pp. 301-9.
- [102] W. Wood and A. Jacinto, "Drosophila melanogaster embryonic haemocytes: masters of multitasking," *Nature Reviews Molecular Cell Biology*, vol. 8, Jul. 2007, pp. 542-51.
- [103] U. Tepass, L.I. Fessler, A. Aziz, and V. Hartenstein, "Embryonic origin of hemocytes and their relationship to cell death in Drosophila," *Development*, vol. 120, Jul. 1994, pp. 1829-37.
- [104] N.K. Cho, L. Keyes, E. Johnson, J. Heller, L. Ryner, F. Karim, and M.A. Krasnow, "Developmental Control of Blood Cell Migration by the Drosophila VEGF Pathway," *Cell*, vol. Vol 108, Mar. 2002, pp. 865-876.
- [105] B. Olofsson and D.T. Page, "Condensation of the central nervous system in embryonic Drosophila is inhibited by blocking hemocyte migration or neural activity," *Developmental Biology*, vol. 279, Mar. 2005, pp. 233-43.
- [106] H.C. Sears, C.J. Kennedy, and P.A. Garrity, "Macrophage-mediated corpse engulfment is required for normal Drosophila CNS morphogenesis," *Development*, vol. 130, Aug. 2003, pp. 3557-3565.
- [107] P. Bangs, N. Franc, and K. White, "Molecular mechanisms of cell death and phagocytosis in Drosophila," *Cell Death and Differentiation*, vol. 7, Nov. 2000, pp. 1027-34.
- [108] N.C. Franc, J.L. Dimarcq, M. Lagueux, J. Hoffmann, and R.A. Ezekowitz, "Croquemort, a novel Drosophila hemocyte/macrophage receptor that recognizes apoptotic cells," *Immunity*, vol. 4, May. 1996, pp. 431-43.
- [109] J.M. Abrams, K. White, L.I. Fessler, and H. Steller, "Programmed cell death during Drosophila embryogenesis," *Development*, vol. 117, Jan. 1993, pp. 29-43.

- [110] B. Stramer, W. Wood, M.J. Galko, M.J. Redd, A. Jacinto, S.M. Parkhurst, and P. Martin, "Live imaging of wound inflammation in *Drosophila* embryos reveals key roles for small GTPases during in vivo cell migration," *Journal of Cell Biology*, vol. 168, Feb. 2005, pp. 567-573.
- [111] W. Wood, C. Faria, and A. Jacinto, "Distinct mechanisms regulate hemocyte chemotaxis during development and wound healing in *Drosophila melanogaster*," *Journal of Cell Biology*, vol. 173, May. 2006, pp. 405-416.
- [112] T.O. Tingvall, E. Roos, and Y. Engström, "The GATA factor Serpent is required for the onset of the humoral immune response in *Drosophila* embryos," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, Mar. 2001, pp. 3884-8.
- [113] J.B. Duffy, "GAL4 system in *Drosophila*: a fly geneticist's Swiss army knife," *Genesis*, New York: Wiley-Liss, 2002, pp. 1-15.
- [114] J. Guan, *Cell Migration: Developmental Methods and Protocols*, Humana Press, 2004.
- [115] J. Colombelli, S.W. Grill, and E.H.K. Stelzer, "Ultraviolet diffraction limited nanosurgery of live biological tissues," *Review of Scientific Instruments*, vol. 75, Feb. 2004, pp. 472-478.
- [116] A. Vogel and V. Venugopalan, "Mechanisms of pulsed laser ablation of biological tissues," *Chemical Reviews*, vol. 103, Feb. 2003, pp. 577-644.
- [117] S.W. Grill, P. Gönczy, E.H. Stelzer, and A.A. Hyman, "Polarity controls forces governing asymmetric spindle positioning in the *Caenorhabditis elegans* embryo," *Nature*, vol. 409, Feb. 2001, pp. 630-3.
- [118] J. Colombelli, E.G. Reynaud, J. Rietdorf, R. Pepperkok, and E.H.K. Stelzer, "In vivo selective cytoskeleton dynamics quantification in interphase cells induced by pulsed ultraviolet laser nanosurgery," *Traffic*, vol. 6, Dec. 2005, pp. 1093-102.
- [119] R. Skaer and S. Whytock, "Interpretation of the three-dimensional structure of living nuclei by specimen tilt," *Journal of Cell Science*, vol. 19, Oct. 1975, pp. 1-10.
- [120] P.J. Shaw, D.A. Agard, Y. Hiraoka, and J.W. Sedat, "Tilted view reconstruction in optical microscopy. Three-dimensional reconstruction of *Drosophila melanogaster* embryo nuclei," *Biophysical Journal*, vol. 55, Jan. 1989, pp. 101-110.
- [121] J. Bradl, M. Hausmann, V. Ehemann, D. Komitowski, and C. Cremer, "A tilting device for three-dimensional microscopy: application to in situ imaging of interphase cell nuclei," *Journal of microscopy*, vol. 168, Oct. 1992, pp. 47-57.
- [122] C.J. Cogswell, K.G. Larkin, and H.U. Klemm, "Fluorescence microtomography: multiangle image acquisition and 3D digital reconstruction," *Three-Dimensional Microscopy: Image Acquisition and Processing III*, 1996, pp. 109-115.
- [123] S. Kikuchi, K. Sonobe, L.S. Sidharta, and N. Ohya, "Three-dimensional computed tomography for optical microscopes," *Optics Communications*, vol. 107, May. 1994, pp. 432-444.
- [124] S. Kikuchi, K. Sonobe, D. Shinohara, N. Ohya, S. Mashiko, and Y. Hiraoka, "A double-axis microscope and its three-dimensional image position adjustment based on an optical marker method," *Optics Communications*, vol. 129, Aug. 1996, pp. 237-244.
- [125] S. Kikuchi, K. Sonobe, S. Mashiko, Y. Hiraoka, and N. Ohya, "Three-dimensional image reconstruction for biological micro-specimens using a double-axis fluorescence microscope," *Optics Communications*, vol. 138, May. 1997, pp. 21-26.
- [126] J. Bradl, M. Hausmann, B. Schneider, B. Rinke, and C. Cremer, "A Versatile 2Pi-Tilting Device for Fluorescence Microscopes," *Journal of Microscopy*, vol. 176, 1994, pp. 211-221.

- [127] M. Kozubek, P. Matula, H. Eipel, and M. Hausmann, "Automated multi-view 3D image acquisition in human genome research," *Proceedings of the First International Symposium on 3D Data Processing Visualization and Transmission.*, 2002, pp. 91-98.
- [128] P.J. Verveer, J. Swoger, F. Pampaloni, K. Greger, M. Marcello, and E.H.K. Stelzer, "High-resolution three-dimensional imaging of large specimens with light sheet-based microscopy," *Nature Methods*, vol. 4, Apr. 2007, pp. 311-313.
- [129] S. Preibisch, T. Rohlfing, M.P. Hasak, and P. Tomancak, "Mosaicing of Single Plane Illumination Microscopy Images Using Groupwise Registration and Fast Content-Based Image Fusion," *Proceedings of SPIE*, vol. 6914, 2008.
- [130] A. Can, O. Al-Kofahi, S. Lasek, D.H. Szarowski, J.N. Turner, and B. Roysam, "Attenuation Correction in Confocal Laser Microscopes: a Novel Two-View Approach," *Journal of Microscopy*, vol. 211, Jul. 2003, pp. 67-79.
- [131] P.J. Shaw, "Three-dimensional optical microscopy using tilted views," *Journal of microscopy*, vol. 158, May. 1990, pp. 165-72.
- [132] M. Kozubek, M. Skalnikova, P. Matula, E. Bartova, J. Rauch, F. Neuhaus, H. Eipel, and M. Hausmann, "Automated microaxial tomography of cell nuclei after specific labelling by fluorescence in situ hybridisation," *Micron*, vol. 33, 2002, pp. 655-665.
- [133] J. Swoger, J. Huisken, and E.H.K. Stelzer, "Multiple imaging axis microscopy improves resolution for thick-sample applications," *Optics Letters*, vol. 28, 2003, pp. 1654-1656.
- [134] A.C. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging*, Society for Industrial Mathematics, 2001.
- [135] W.C. Röntgen, "On a new kind of rays," *Science*, vol. 3, Feb. 1896, pp. 227-231.
- [136] G.N. Hounsfield, "Method and apparatus for measuring X- or gamma-radiation absorption or transmission at plural angles and analyzing the data," U.S. Patent 3778614, December 1973.
- [137] R.G. Hart, "Electron Microscopy of Unstained Biological Material: The Polytropic Montage," *Science*, vol. 159, Mar. 1968, pp. 1464-1467.
- [138] D.J. De Rosier and A. Klug, "Reconstruction of Three Dimensional Structures from Electron Micrographs," *Nature*, vol. 217, Jan. 1968, pp. 130-134.
- [139] A.J. Koster, R. Grimm, D. Typke, R. Hegerl, A. Stoschek, J. Walz, and W. Baumeister, "Perspectives of molecular and cellular electron tomography," *Journal of Structural Biology*, vol. 120, Dec. 1997, pp. 276-308.
- [140] V. Lucić, F. Förster, and W. Baumeister, "Structural studies by electron tomography: from cells to molecules," *Annual Review of Biochemistry*, vol. 74, 2005, pp. 833-65.
- [141] K.H. Downing, H. Sui, and M. Auer, "Electron tomography: A 3D View of the Subcellular World," *Analytical Chemistry*, vol. 79, Nov. 2001, pp. 7949-57.
- [142] J. Sharpe, U. Ahlgren, P. Perry, B. Hill, A. Ross, J. Hecksher-Sorensen, R. Baldock, and D. Davidson, "Optical Projection Tomography as a Tool for 3D Microscopy and Gene Expression Studies," *Science*, vol. 296, Apr. 2002, pp. 541-545.
- [143] W. Wein, M. Blume, U. Leischner, H. Dodt, and N. Navab, "Quality-Based Registration and Reconstruction of Optical Tomography Volumes," *Proceedings of MICCAI*, 2007, pp. 718-725.
- [144] A.K. Jain, *Fundamentals of Digital Image Processing*, Prentice Hall, 1988.
- [145] L.G. Brown, "A survey of image registration techniques," *ACM Computing Surveys*, vol. 24, 1992, pp. 325-376.
- [146] J.B.A. Maintz and M.A. Viergever, "A survey of medical image registration," *Medical Image Analysis*, vol. 2, Mar. 1998.

- [147] B. Zitova and J. Flusser, "Image registration methods: a survey," *Image and Vision Computing*, vol. 21, Oct. 2003, pp. 977-1000.
- [148] A. Rosenfeld and A.C. Kak, *Digital Picture Processing*, Academic Press, 1982.
- [149] R.O. Duda and P.E. Hart, *Pattern classification and scene analysis*, New York: Wiley, 1973.
- [150] W.M.I. Wells, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis, "Multi-modal volume registration by maximization of mutual information," *Medical Image Analysis*, vol. 1, Mar. 1996, pp. 35-51.
- [151] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE Transactions on Medical Imaging*, vol. 16, 1997, pp. 187-198.
- [152] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery, *Numerical Recipes: The Art of Scientific Computing*, Cambridge University Press, 2007.
- [153] E.D. Castro and C. Morandi, "Registration of translated and rotated images using finite Fourier transforms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 9, 1987, pp. 700-703.
- [154] P. Viola and W.M. Wells III, "Alignment by Maximization of Mutual Information," *International Journal of Computer Vision*, vol. 24, 1997, pp. 137-154.
- [155] T.M. Cover and J.A. Thomas, *Elements of Information Theory*, Wiley & Sons, 2006.
- [156] I. Vajda, *Theory of Statistical Inference and Information*, Springer, 1989.
- [157] S. Preibisch, T. Rohlfing, M.P. Hasak, and P. Tomancak, "Mosaicing of single-plane illumination microscopy images using groupwise registration and fast content-based image fusion," *Proceedings of SPIE*, vol. 6914, 2008.
- [158] R.M. Haralick and L.G. Shapiro, *Computer and Robot Vision, Vol. 1*, Boston: Addison-Wesley, 1992.
- [159] U. Kržič and E.H.K. Stelzer, "Real-space based multi-view image fusion," Conference presentation, Focus on microscopy 2007, Valencia, Apr. 2007.
- [160] P. Soille, *Morphological Image Analysis*, Springer, 2004.
- [161] A.A. Goshtasby, "Fusion of multi-exposure images," *Image and Vision Computing*, vol. 23, Jun. 2005, pp. 611-618.
- [162] S.A. Prahl, M. Keijzer, S.L. Jacques, and A.J. Welch, "A Monte Carlo model of light propagation in tissue," *SPIE Institute Series*, vol. 5, 1989, pp. 102-1.
- [163] L.B. Lucy, "An iterative technique for the rectification of observed distributions," *The Astronomical Journal*, vol. 79, 1974, p. 745.
- [164] K. Satzler and R. Eils, "Resolution improvement by 3-D reconstructions from tilted views in axial tomography and confocal theta microscopy," *Bioimaging*, vol. 5, 1997, pp. 171-182.
- [165] N. Wiener, *Extrapolation, Interpolation, and Smoothing of Stationary Time Series with Engineering Applications Application*, The MIT Press, 1949.
- [166] R.C. Gonzalez, R.E. Woods, and S.L. Eddins, *Digital Image Processing Using MATLAB*, Prentice Hall, 2003.
- [167] Jan Huisken, "Multiview microscopy and multibeam manipulation for high resolution optical imaging," Ph.D. thesis, Albert-Ludwigs-Universität Freiburg im Breisgau, 2004.
- [168] P.J. Verveer, M.J. Gemkow, and T.M. Jovin, "A comparison of image restoration approaches applied to three-dimensional confocal and wide-field fluorescence microscopy," *Journal of Microscopy*, vol. 193, 1999, pp. 50-61.

- [169] P.J. Verveer and T.M. Jovin, “Improved restoration from multiple images of a single object: application to fluorescence microscopy,” *Applied Optics*, vol. 37, 1998, pp. 6240–6246.
- [170] W.H. Richardson, “Bayesian-based iterative method of image restoration,” *Journal of the Optical Society of America*, vol. 62, 1972, pp. 55-59.
- [171] M.K. Singh, U.S. Tiwary, and Y. Kim, “An adaptively accelerated Lucy-Richardson method for image deblurring,” *EURASIP Journal on Advances in Signal Processing*, vol. 8, 2008.
- [172] H. Volkmann, “Ernst Abbe and His Work,” *Applied Optics*, vol. 5, Nov. 1966, pp. 1720-1731.
- [173] J.R. Lakowicz, *Principles of Fluorescence Spectroscopy*, Springer, 2006.
- [174] 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*, vol. 24, 1981, pp. 381-395.

ABBREVIATIONS

2D/3D	two-Dimensional, three-Dimensional
ADC	Analog-to-Digital Converter
AOTF	Acousto-Optical Tunable Filter
CARS	Coherent Anti-Stokes Raman Scattering
CCD	Charge-Coupled Device
CLSM	Confocal Laser Scanning Microscope
CSOM	Computational Optical Sectioning Microscopy
CT	Computed Tomography
DAPI	4',6-Diamidino-2-Phenylindole
DSLM	Digital Scanned laser Light-sheet Microscope
FFT	Fast Fourier Transform
FOV	Field Of View
FRAP	Fluorescence Recovery After Photo-bleaching
FWD	Free Working Distance (of an objective lens)
GFP	Green Fluorescent Protein
LSFM	Light-Sheet based Fluorescence Microscope
LST	Light-Sheet Thickness
MDCK	Madin-Darby Canine Kidney (cell line)
MIAM	Multiple Imaging Axis Microscope
MTF	Modulation Transfer Function
NA	Numerical Aperture
NCC	Normalized Cross-Correlation
NMR	Nuclear Magnetic Resonance
OPT	Optical Projection Tomography
PALM	Photo-Activated Localization Microscopy
PMT	Photo-Multiplier Tube
PSF	Point Spread Function
RFP	Red Fluorescent Protein
SGH	Stelzer-Grill-Heisenberg (theory)
SNOM	Scanning Near-field Optical Microscope
SNR	Signal-to-Noise Ratio
SPIM	Single Plane Illumination Microscope
SPIM-SI	SPIM with Structured Illumination
STED	Stimulated Emission Depletion
STORM	STochastic Optical Reconstruction Microscopy
SWFM	Standing Wave Fluorescence Microscope
TEM	Transmission Electron Microscope
UV	Ultra-Violet

TABLE OF FIGURES

Figure 1: Jablonski diagram of a typical fluorophore.....	9
Figure 2: Examples of fluorescent protein spectra.	10
Figure 3: Diagram of an epifluorescence microscope.	15
Figure 4: Resolution of a wide-field microscope.	16
Figure 5: Intensity point spread function of a wide-field microscope.....	18
Figure 6: Diagram of a confocal microscope.	20
Figure 7: Intensity point spread function of a confocal microscope.	22
Figure 8: Diagram of a multi-photon epifluorescence microscope.	23
Figure 9: Diagram of a confocal ϑ -fluorescence microscope.....	25
Figure 10: Intensity PSF of a confocal ϑ -fluorescence microscope.....	26
Figure 11: Illumination and detection in LSFM.....	32
Figure 12: A semi-quantitative comparison of photo-bleaching rates in a SPIM and a regular widefield fluorescence microscope.....	33
Figure 13: Detection arm of a LSFM.	35
Figure 14: Experimental chamber of a LSFM.....	38
Figure 15: Photograph of a LSFM experimental chamber.	38
Figure 16: Light-sheet dimensions.	41
Figure 17: Light-sheet height.	43
Figure 18: The axial resolution of a LSFM.	45
Figure 19: Intensity PSF of a LSFM for three common detection objective lenses.....	46
Figure 20: Contrast of the LSFM.	47
Figure 21: Two types of SPIM illumination.....	48
Figure 22: Light-sheet thickness.	49
Figure 23: Basic SPIM setup.....	50
Figure 24: DSLM illumination.....	51
Figure 25: LSFM translation/rotation stage.	52

Figure 26: Common ways of specimen mounting in LSFM.....	53
Figure 27: Effect of agarose depth on image quality.....	54
Figure 28: Effect of agarose concentration on image quality.....	55
Figure 29: Effect of agarose concentration on bead mobility.....	56
Figure 30: Effect of agarose concentration on bead mobility.....	56
Figure 31: Hemocytes in <i>Drosophila m.</i> embryo.....	57
Figure 32: Hemocyte migratory pathways in an intact <i>Drosophila m.</i> embryo.....	58
Figure 33: Automated Hemocyte detection in SPIM images.....	60
Figure 34: SPIM laser-cutter optical setup.....	61
Figure 35: Hemocyte response to a UV laser induced wound.....	62
Figure 36: Hemocyte response to a UV laser induced wound.....	63
Figure 37: Resolution improvement by multiple-view imaging.....	67
Figure 38: Multiple-view image fusion with opaque specimen.....	69
Figure 39: Rotation calibration for LSFM multiple-view imaging.....	73
Figure 40: Center of rotation error translation.....	75
Figure 41: Demonstration of centred rotation.....	76
Figure 42: Main parts of image fusion algorithm.....	78
Figure 43: An example of the pair-wise alignment of multiple-views.....	81
Figure 44: Arbitrary vector and its eight nearest grid points.....	85
Figure 45: An example of image artifacts due to inaccurate image alignment.....	87
Figure 46: Tilted axis of rotation.....	88
Figure 47: Sources of skew in LSFM images.....	90
Figure 48: Speed improvement by pyramid alignment of LSFM views.....	92
Figure 49: Effect of image deformation correction on image alignment.....	96
Figure 50: Average and maximum intensity fusion methods in case of a non-transparent specimen.....	97
Figure 51: Average and maximum intensity fusion methods in case of a semi-transparent specimen.....	98
Figure 52: Derivative weighted fusion method.....	101
Figure 53: Image amplitude spectrum envelope and Lucy-Richardson deconvolution.....	112
Figure 54: Image amplitude spectrum envelope and iterative expectation maximization fusion algorithm.....	113
Figure 55: Attainable spectrum coverage as a function of number of fusion iterations.....	115
Figure 56: Convergence speed as a function of the number of views.....	117

Figure 57: Fusion of images of point-object.	118
Figure 58: Size of a point-object after a number of two-view fusion iterations.....	119
Figure 59: Resolution improvement due to fusion/deconvolution.	120
Figure 60: Resolution improvement due to fusion/deconvolution.	121
Figure 61: Fusion of simulated multiple-view images.	122
Figure 62: Fusion of simulated multiple-view images.	123
Figure 63: Image restoration by multiple-view fusion.	124
Figure 64: Multiple views of <i>Saccharomyces cerevisiae</i> expressing FIL1-GFP.....	125
Figure 65: Multiple views of <i>Saccharomyces cerevisiae</i> labeled with Phalloidin-GFP (marks actin cables).....	126
Figure 66: Comparison of different fusion techniques – transparent specimen.	127
Figure 67: Comparison of different fusion techniques – semi-transparent specimen.	128
Figure 68: Input images and their fusion.....	129
Figure 69: Multiple-view image of a fully opaque specimen.....	130
Figure 70: Multiple-view image of a large and opaque specimen’s surface.	131
Figure 71: Single view and multiple-views fusion of juvenile Medaka fish with an acetylated tubulin immuno-staining.....	132

ACKNOWLEDGEMENTS

The work presented in this thesis has been carried out in the Light Microscopy Group of the European Molecular Biology Laboratory (EMBL) in Heidelberg. I am grateful to Ernst Stelzer for giving me a chance to work in his research group, for introducing me to the modern microscopy and for supervising the progress of my research project. Many thanks also to Klaus Greger for helping me with the first steps in the world of SPIM, to Emmanuel Reynaud for personal support and for being the “biologist on duty” whenever (and I mean *whenever*) help was needed, and to Julien Colombelli, whose original opinions were always a priceless source of inspiration. You were more than just colleagues. Thanks also to all the other former and present members of the light microscopy group that have helped generating a friendly and motivating research environment. Jim Swoger, Fabian Härle, Francesco Pampaloni, Khaled Khairy, Petra Jakob, Holger Kreß, David Dickson, Chrisoph Engelbrecht, Manuel Neetz, Philippe Girard, Tobias Breuninger, Jan Huisken, Philipp Keller and Alexander Rohrbach, it was a pleasure to work with you.

I would like to thank Prof. Jörg Schmiedmayer for being an external supervisor of my project and for many helpful conversations. Thanks also to other members of my thesis advisory committee: Darren Gilmour, Detlev Arendt and Pernille Rørth. Your ideas and recommendations are much appreciated.

Development of new instruments would have been impossible without the help of Leo Burger, Helmut Schaar, Alfons Riedinger, Georg Ritter and all other members of the EMBL’s electronics and mechanical workshops.

My doctoral work was financially supported by the Fondation Louis-Jeantet from Geneva.

Special thanks to Belinda Bullard, Kevin Leonard and Igor Mušević for introducing me to the world of science.

I would also like to thank my dear friends Sini, David, Manu, Klaus, Ervin, Julien, Daniela, Fabian and Barbara, Sonja, Eneja, Andrej, Boštjan, Peter, Tomaž, Aksel, Helena, Vesna, Iva, Barbara, Hendrik, Rebecca, Jenny, Florian, Lucija, Stina, Vicente, Anika, members of EMBL’s dive club and all the others who kept reminding me that there is more to life than science. I hope we will keep in touch for many years to come despite the distances between us.

Finally, I would like to thank my father Andrej, mother Katja, and my brothers Jernej and Andrej for the kind of support that only a loving family can offer.

And last but not least, Alenka, thank you for bringing the sun into my life.