UNIVERSITY OF KENT

DOCTORAL THESIS

Computational Optical Sectioning in Fibre Bundle Endomicroscopy

Author:
Andrew David
THRAPP

Supervisor:
Dr. Michael R.
HUGHES

A thesis submitted in fulfilment of the requirements
for the degree of Doctor of Philosophy

in the

Applied Optics Group
School of Physical Sciences

December 2020

Declaration of Authorship

I, Andrew D. THRAPP, declare that this thesis titled, "Computational Optical Sectioning in Fibre Bundle Endomicroscopy" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed: *Andrew Thrapp*

Date: 2 Dec 2020

Last compiled: Sunday 28th March, 2021 at 10:29pm.

"Ideas do not always come in a flash but by diligent trial-and-error experiments that take time and thought. "

Charles K. Kao, 2009 Nobel Prize in Physics

UNIVERSITY OF KENT

Abstract

Faculty of Sciences
School of Physical Sciences

Doctor of Philosophy

Computational Optical Sectioning in Fibre Bundle Endomicroscopy

by Andrew D. THRAPP

THE field of fibre bundle endomicroscopy has emerged to enable real-time imaging of cellular level features *in-vivo*. The gold standard is confocal laser scanning, enabling optical sectioning. Point-scan confocal suffers from lower speeds, a need for complex alignment, and the added cost of a laser. This thesis presents three developments in computational optical sectioning for fibre bundle endomicroscopy.

The first development is in structured illumination (SIM) endomicroscopy. Lower-cost, simplified endomicroscopes have been developed which use wide-field incoherent illumination. Optical sectioning can be introduced to these systems using SIM. SIM improves imaging using spatial modulation of the focal plane and capturing a three-frame sequence. The acquired images are then numerically processed to reject out-of-focus light. This thesis reports and characterises the first high-speed SIM endomicroscope built using a miniature array of mirrors, a digital micromirror device.

The second development is automated motion compensation in SIM endomicroscopy. As a multi frame process, SIM is susceptible to motion artefacts, making the technique difficult to use *in vivo* and preventing the use of mosaicking to synthesise a larger effective field of view. I report and validate an automatic motion compensation technique to overcome motion artefacts and report the first mosaics in SIM endomicroscopy.

The third development is improvements in subtraction-based enhanced line scanning (ELS) endomicroscopy. The 2D scanning of a point scan confocal endomicroscope can be replaced by a scanning line which is synchronised to the sequential readout of a rolling shutter camera. While this leads to high-speed sectioning, as with all line scanning systems, far-from-focus light degrades images. It is possible to remove this by subtracting a second image taken with an offset detection slit. This has previously required two-cameras or two sequential frames. The latter introduces motion artefacts. This thesis presents a novel approach to ELS using single frame acquisition with real-time mosaicking at 240 frames/s.

Acknowledgements

THIS thesis has only been possible due to a large number of people who I have been lucky enough to come into contact with and get to know over the last few years.

Throughout the writing of this dissertation, I have received a great deal of support and assistance. I would first like to thank my main supervisor Dr Mike Hughes. I am very grateful for Mike's expertise, and willingness to constantly work with me to develop my ideas.

I would like to acknowledge my colleagues at the Applied Optics Group (AOG) for their wonderful support and collaboration. AOG has built a very supportive community, of people who are more than willing to work together for shared success. I would like to thank my second supervisor Prof Adrian Podoleanu and Dr Manuel Marques, who have both always been willing to offer expertise and helpful advice. Starting with my fellow students, some of whom have since moved on, I would like to thank Adrian Fernandez-Uceda, Gianni Nteroli, Julien Camard, Rene Riha, Alejandro Marinez Jiimenez, Melanie Whacker, Dr Sophie Cajoulle, Mike Everson, and many other exchange students and external researchers for our many board game nights, movie nights, and office conversations.

Also, I would like to thank my family who has always inspired me. Including my father, whose work ethic and moral compass always drove me to be better, my mother, whose intellectual curiosity led me to question everything, my grandparents, who have always been a source of unconditional support, my stepfather, who helped me see the value in entrepreneurship and the good in humanity, my sisters Lynn, and Jacki, who are both wonderful people, and have accomplished a lot in their own right. Last but certainly not least, my brother Conner, who hasn't chosen his path yet but to who I wish all the happiness success in the world. If I contribute anything to the world it is built on the foundations these people have provided me with.

I would like to thank the University of Kent for the Vice-Chancellor's Graduate Teaching Assistantship, as well as the International School of Light and Science Technologies (ISLiST) for being awarded a scholarship for two years to attend their summer school, the Optical Society of America for providing numerous travel grants to present work in Nottingham, and Barcelona. As well as Guy Holmes with Laser 2000 for working with our OSA chapter to build stronger links to industry and coordinating a science fair in Huntingdon. I would also like to thank Gemma Hunt, Dr Vicky Mason, and Hannah Thompson for the outreach opportunities.

Publications

The publications that resulted from this thesis are listed below:

A. Thrapp and M. Hughes, "Automatic motion compensation for structured illumination endomicroscopy using a flexible fiber bundle," *J Biomed Opt*, vol. 25, no. 2, pp. 1–13, 2020, ISSN: 1560-2281 (Electronic) 1083-3668 (Linking). DOI: 10.1117/1.JBO.25.2.026501.

M. J. Marques, M. R. Hughes, K. Vyas, **A. Thrapp**, H. Zhang, A. Bradu, G. Gelikonov, P. Giataganas, C. J. Payne, G. Z. Yang, and A. Podoleanu, "En-face optical coherence tomography/fluorescence endomicroscopy for minimally invasive imaging using a robotic scanner," *J Biomed Opt*, vol. 24, no. 6, pp. 1–15, 2019, ISSN: 1560-2281 (Electronic) 1083-3668 (Linking). DOI: 10.1117/1.JBO.24.6.066006.

A. Thrapp, M. R. Hughes (2019, February). "Motion compensation in structured illumination fluorescence endomicroscopy." In *Endoscopic Microscopy XIV* (Vol. 10854, p. 108541D). International Society for Optics and Photonics.Proc. DOI: 10.1117/12.2509590

A. Thrapp and M. Hughes, "Reduced motion artifacts and speed improvements in enhanced line scanning fiber bundle endomicroscopy." *J Biomed Opt*. Accepted Mar, 2021, publication date pending.

Contents

List of Abbreviations

BS	**B**eam**s**plitter
CCD	**C**harge **C**oupled **D**evice
CL	**C**ondenser **L**ens
CMOS	**C**omplimentary **M**etal **O**xide **S**emiconductor
DBS	**D**ichroic **B**eam**s**plitter
DET	**Det**ector
DMD	**D**igital **M**icromirror **D**evice
F	**F**ilter
FB	**F**ibre **B**undle
FWHM	**F**ull **W**idth at **H**alf **M**aximum
GS	**G**alvanometer **S**canner
HWHM	**H**alf **W**idth at **H**alf **M**aximum
ICS	**I**nter **C**ore **S**pacing
L	**L**ens
LPW	**L**ine **P**air **W**idth
M	**M**irror
MMF	**M**ulti-**M**ode Fibre
OBJ	**Obj**ective lens
PH	**P**in **h**ole
SF	**S**ingle **F**ibre
SIM	**S**tructured **I**llumination **M**icroscopy
SMF	**S**ingle-**M**ode Fibre
TL	**T**ube **L**ens
WF	**W**ide**f**ield

Chapter 1

Introduction

1.1 Tissue imaging

Microscopic tissue imaging can aid clinicians in the diagnosis of cancerous lesions, help diagnose problems at an early stage, and be used as an interoperative tool. One common use of tissue imaging is as a histology. Histology is the process where tissue is removed from the body, frozen, thinly sliced, stained, and then imaged. A high contrast stain such as hematoxylin and eosin (H & E) makes it possible for a clinician to differentiate features in tissue.

The conventional pathology process has a few inherent drawbacks. The multistep process can lead to follow up visits for a patient, bulky optics require adequate facilities to use and store, and multiple specialists are often needed to perform a diagnosis. This process translates to inconvenience and cost to the patient and provider, and time lost in the operating theatre. In rural regions of the world, where infrastructure and specialists may not be available, this can lead to patients not getting into treatment or getting into treatment at a late stage of disease progression [1].

An alternative to the conventional workflow, and gaining clinical relevance, is the optical biopsy via approaches such as endomicroscopy. Using a conventional

endoscope a clinician can investigate accessible and hollow organs such as skin, mouth, oesophagus or GI tract. Endoscopes typically have a large field of view (FOV). This large FOV is useful to get a broad overview of tissue morphology; however, there is a trade-off between the magnification of the lens and the FOV [2]. Structural, as well as cellular level indicators, are clinically relevant indicators [3, 4]. At the cellular level, this includes nuclear crowding, nuclear to cytoplasm ratio, and nucleus size. These are generally not resolvable with an endoscope. A clinical goal of endomicroscopy is to image them. A challenge of optical biopsy is not just to maintain the resolution of conventional microscopy, but to enable the feature discrimination, such as that provided with H & E stained tissue, as well.

1.2 Endomicroscopy

A basic endomicroscope (Fig. 1.1) can be made with a source, detector, proximal optics, distal optics, and a fibre optic waveguide. The source is typically either an LED or a laser, and the detector is either a camera, for full-field imaging, or a point detector such as a photomultiplier tube for scanning systems.

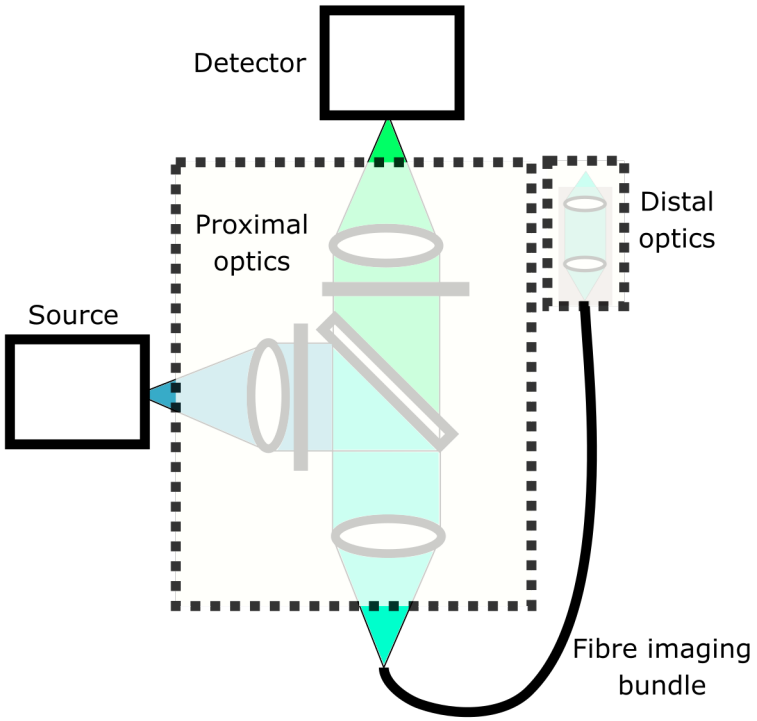


FIGURE 1.1: Core components of a fibre based endomicroscope. Source guides light onto the proximal face of a fibre optic waveguide, light is then either relayed to a sample in contact, or via distal optics and, if applicable, distal scanning elements. Returning light, either reflected, or fluorescence emission is relayed to a detector.

Endomicroscopes can be either full-field (without optical sectioning), scanning or non-scanning. Non-scanning full field operation can be implemented using a configuration like Fig. 1.1 with the waveguide being a multicore ($\sim$ 30,000) fibre optic bundle. A scanning system can be implemented either by placing the multicore fibre bundle at the focal plane of a confocal microscope or by implementing distal scanning with a single or double-clad fibre.

Resolution in conventional microscopy is determined by scalar diffraction theory. Using the Rayleigh criterion lateral resolution of a conventional lens is given by [5]:

$$\delta_{\rho_0} = \frac{0.61\ \lambda}{NA} \tag{1.1}$$

where, NA is the numerical aperture of the lens, and λ is the wavelength of light.

In conventional imaging, resolution can be improved either using a shorter wavelength or by choosing a higher NA lens. The work done in this thesis will use 488 nm light, a fibre imaging bundle with a 3.2 μm inter-core spacing and NA of 0.39. Substitution into Eqn. 1.1, yields a diffraction-limited resolution of 0.78 μm. Due to bundle sampling, it will not be possible to realise this and lateral resolution will always be limited by the inter-core spacing (Sect. 2.2).

In microscopy, contrast can be derived from several sources, including bright-field, darkfield, reflectance, phase gradients, as well as fluorescence.

Brightfield contrast is obtained by back illuminating a sample, and the absorption properties of the sample determine the contrast. While common in benchtop systems, the usefulness of brightfield in endomicroscopy is limited, since in-vivo tissue cannot be thinly sliced or easily back-illuminated.

Reflection endomicroscopy generates contrast based on the wavelength-dependent reflection coefficients of the superficial tissue. Contrast can also come from staining with a compound such as H & E. The sample can be illuminated in an epi-configuration, where the illumination objective is the same as the collection, or in a configuration where the illumination and collection lenses can be decoupled. Reflectance based systems are limited in fibre-based endomicroscopy by reflections from the proximal face of the imaging bundle, discussed in more depth in Sect. 2.2.3. Systems have been developed to overcome back reflections, such as through changing the way light is coupled into the waveguide (Sect. 2.7.3).

Darkfield contrast is obtained by illuminating a sample at an extreme angle so that only light which is scattered by the sample is imaged, and the illumination

light is not. Extreme angles can be generated by blocking the central part of a collimated beam, with a circular stop, in the Fourier plane of an objective. The edge of the objective will then illuminate the sample at an extreme angle, and only scattered light will return to the centre of the lens. As discussed in Sect. 2.1, darkfield illumination has been used to overcome specular reflections from a fibre bundle.

Phase gradient microscopy is a form of transmission imaging capable of resolving refractive index gradients. When light traversing a sample is obliquely incident on a phase gradient, its path deviates generating contrast [6, 7]. Phase-contrast has also been demonstrated in endomicroscopy using a fibre bundle system with illumination from outside the bundle generating an effective source deep in a scattering sample [8].

Fluorescence imaging is a way to derive contrast through fluorescent emission of molecules. It can come from the endogenous tissue compounds such as collagen, or staining agents. In single-photon fluorescence, high energy light excites an electron which relaxes and emits lower-energy light. A commonly used fluorescent staining agent is acridine orange, which binds to RNA and DNA, is used to stain tissue, which provides the contrast [9, 10]. It has been used to image tissue as part of Mohs surgery. Mohs is an approach where suspected cancerous tissue is removed in small slices and examined intraoperatively leading to a reduction in the amount of excised tissue [11].

Fluorescence is the contrast mechanism utilised in the experimental work in this thesis. The compound Acriflavine Hydrochloride (0.05%) is used for system validation with animal tissue. Acriflavine has an excitation wavelength of 463

nm and emission wavelength at 490. Typical concentrations are 0.01-0.05% [12].

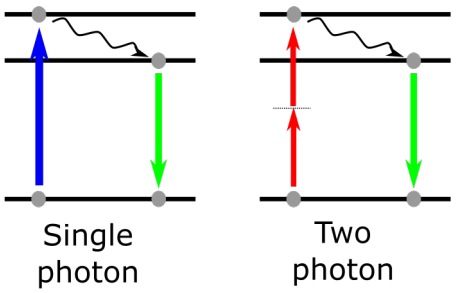


FIGURE 1.2: Principle of single photon fluorescence and two photon fluorescence.

Optical sectioning is critical to reduce imaged degradation caused by out-of-focus features. The effects of blur are apparent when imaging a multiplanar object. Fig 1.3 shows a camera above a staircase imaged with a shallow depth of field. The steps above and below are contributing blur the resulting image. The goal of optical sectioning is to remove that blur, and in the case of fibre bundle endomicroscopy, isolate the focal plane.

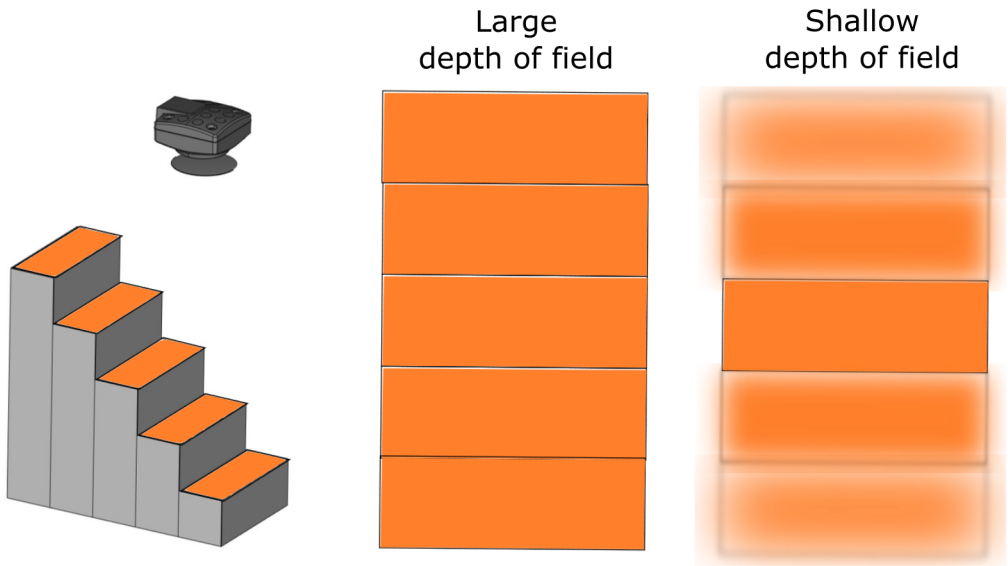


FIGURE 1.3: Multiplanar surface when viewed from a camera above. Large and shallow depth of field shown.

In a clinical setting, a clinician may want to observe a superficial tissue layer or only a thin slice of a thicker sample. To address this, a confocal microscope,

used to perform optical sectioning was developed by Minsky, M. [13] (Fig. 1.4). Confocal microscopy uses a point source and a pinhole to block light which does not emanate from focus.

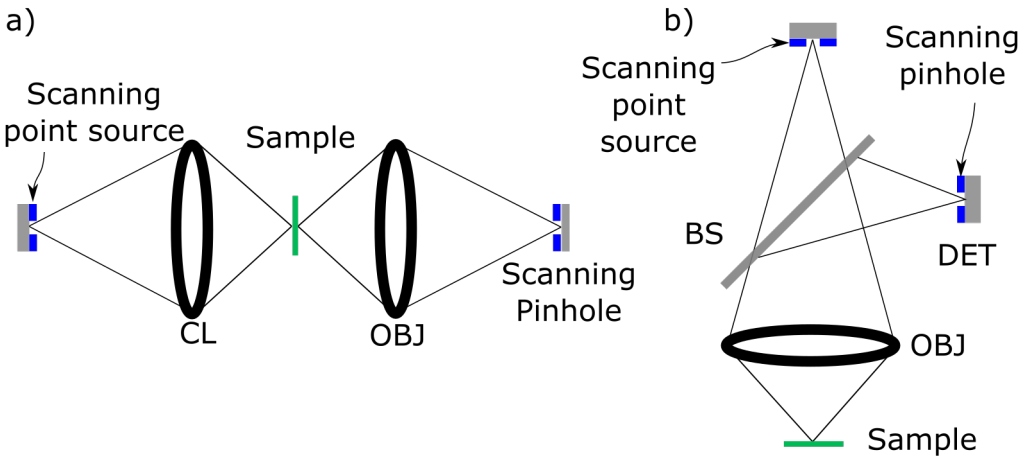


FIGURE 1.4: Confocal microscope with a point source and pinhole detector in (a) transmission, and (b) reflectance configuration. In the configurations shown, scanning between the pinhole and point-source must be synchronised.

Imaging in an endomicroscope is typically done using either a single fibre optic probe, a fibre optic bundle, or side-viewing probe. The distal tips of example probes are shown in Fig. 1.5. There are trade-offs between lensed systems and lensless systems; distal optics allow focusing, but the rigid tip makes it harder to traverse the working channel of an endoscope.

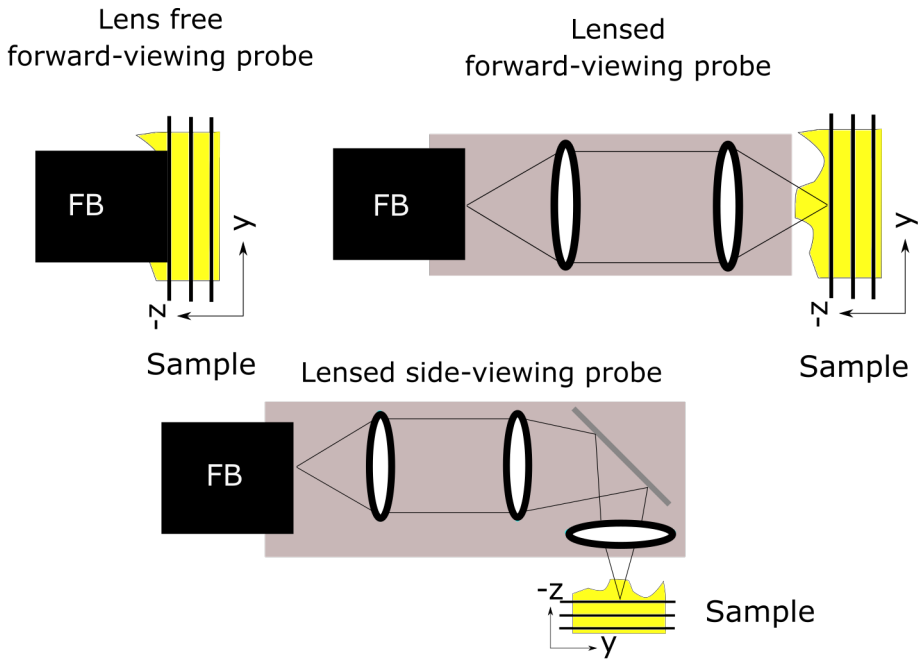


FIGURE 1.5: Fibre bundle distal probe tips, a multicore fibre optic bundle is the core of all three approaches, lens-free, and lensed forward and side-viewing. Based on designs described in Ref. [2, 85].

Typical resolutions for endomicroscopes are around 5 μm with shallow depth penetration, and a typical a field of view (FOV) is under 1 mm. There is a trade-off between resolution and field of view. A lensed probe typically has higher resolution with a smaller field of view. The speed can vary vastly with commercial systems reporting around 1-15 fps, and research systems reporting at a video rate of around 15-120 Hz.

Mosaicking is a useful way to generate a larger field-of-view without a larger, and more invasive, probe. Mosaicking works by stitching together sequentially acquired images. Mosaicking can be done in real-time, using methods such as rigid registration and dead leaves image insertion [14]. Mosaicking can also be done offline, allowing for tissue deformation correction, correction of the accumulated registration error, and, in some cases, improve resolution [15–17]. This is discussed in more depth in Sect. 3.1.

1.3 Clinical applications

Endomicroscopy is driven by clinical demand. Imaging can be performed on accessible organs such as the skin [18], mouth [19], or cervix [20]. It can also be performed in endoscopically accessible organs such as the oesophagus [21], gastrointestinal (GI) tract [22], or lung [23], or surgically accessible organs such as the brain [24], prostate [25], ovaries [26], and breasts [27].

A 2005 report evaluated the feasibility of confocal laser endomicroscopy (CLE) for GI tract imaging. The report concluded that CLE was capable of resolving cellular structures as well as subcellular such as nuclei. The study also reported that the contrast agent Acriflavine strongly contrasted the cell nuclei of the surface epithelium [28]. Mapping studies (2007) of the small and large intestine from 45 patients have also been performed demonstrating good concordance between histological appearances and CLE images [29]. A review of CLE in gut pathology is available in Ref. [29]

CLE has also been demonstrated for reported for *ex-vivo* analysis of freshly excised tissue during breast-conserving surgery (2015) [30]. The accuracy of 17 histopathologists and surgeons, who had not used confocal laser endomicroscopy systems before, to reproduce histological findings was 94% and 92% respectively. Inter-operative agreement was "almost perfect" for histopathologists and "very good" for surgeons. A different CLE pilot study for breast-conserving surgery also reported CLE images with architectural features that corresponded to histopathological images [31].

In 2014 Liu, J. et al., demonstrated brain imaging using CLE [32] and reported

several image-guided approaches to intraoperatively determine margins, leading to a reduction in unnecessary tissue removal. Multiple dyes were used at multiple wavelengths. This *in vivo* assessment used a Zeiss Optiscan (Sect. 2.4) [33] as the CLE, which was limited to 0.8 frames per second.

One case study is using endomicroscopy in the diagnosis of melanocytic skin tumours [34]. Depending on the expertise of a dermatologist the sensitivity, true positive rate, was on average 70.1% in a range between 61-91% with the 91% being reported in a single study with dermatologists who had ten or more years' experience, and a specificity, true negative rate, of 99.4%. A 2005 study showed that observers with no previous instruction in CLE, achieved a sensitivity of 88.15% and specificity of 97.6% [35]. CLE delivered a 25% improvement in selectivity over four-quadrant endoscopic biopsy alone.

The high-resolution microendoscope (HRME) is a lower cost, full-field endomicroscope (Sect. 2.1). It has been demonstrated (2010) as part of an interoperative screening of Barrett's oesophagus, a precursor to oesophagal adenocarcinoma, a particularly fatal condition, with a five-year survival rate of 14% and 0% for stage III and stage IV, respectively [36]. It is characterised by tissue in the oesophagus beginning to resemble tissue in the stomach. White light endoscopy is shown to miss abnormal growths in 61% of cases [36]. Muldoon et al., with the HRME reported 87% sensitivity and 61% specificity. A 2015 study presented a quantitative analysis of HRME images of Oesophageal tissue [37].

Low resource settings experience more than 80% of the total incidence of cervical cancer and mortality [4, 38]. Cervical intraepithelial neoplasia is characterised by an increase in nuclear density, size, and atypical cell shapes [4]. HRME

reported 97% sensitivity and specificity of 80% for detecting grade 1, and 93% for grades 2, and 3. HRME has been deployed to rural Botswana for on-the-spot diagnostics [38].

1.4 Contributions of thesis

Low-cost systems such as those developed with LED illumination are helping bring endomicroscopy into the clinical domain. Additionally, they are putting emerging clinical modalities into regions of the world where facilities and specialists are not available [39]. The HRME (Sect. 2.1) has shown promising results for in-vivo imaging [40, 41]. HRME, however, has limited inherent optical sectioning meaning images are degraded by blur. Optical sectioning can be added with structured illumination (Sect. 4). SIM is a multi-frame technique involving three images where high spatial frequency square wave modulation is imaged onto the focal plane. There is a trade-off relationship between optical sectioning and signal, and on-the-fly pattern customisation would allow a clinician to balance this in real time. Nuclear morphometry is quantitative analysis of changes in nuclear size and shape. A structured illumination HRME has been reported to obtain nuclear morphometry in columnar epithelium such as oral mucosa and the cervix [42] [43].

A SIM endomicroscope using a digital micromirror device (DMD) has not been previously implemented [42]. A DMD is a high-speed array of mirrors.

The DMD allows for tuning of the modulation frequency of the structured illumination to balance the signal and optical sectioning strength in real-time. An approach to reduce DMD grid discretisation effects, as well as a parameter optimisation is presented (optical transfer function, axial response, modulation depth). SIM images of lens paper are reconstructed, showing expected background rejection when using different modulation frequencies. The approach is then validated with porcine stomach, showing blur reduction in the thick scattering tissue. Furthermore, it has been reported that images can be depth-resolved [44]. By taking images at three different modulation spatial frequencies, a subtraction based approach is used to show that this can be done using a DMD. The DMD can also be used in full-field mode, making it effectively a HRME. The system constructed in this thesis (Sect. 4), operated in this mode, was contributed to a multimodal endoscopic OCT / endomicroscopy collaboration using a robotic arm. This contribution, which led to publication (Ref. [45]), reported combined OCT and endomicroscopy mosaics.

Fibre probes typically have a diameter of under 1 mm, and while this means they can maintain endoscopic compatibility, the FOV is often much smaller than typical histology. It can be useful to stitch together adjacent frames into a mosaic. Problematically, SIM is a multi-frame process, and even slight shifts between raw frames can lead to artefacts. These motion artefacts have meant that, before the work presented in this thesis, which led to publication [17, 46], mosaics have not been previously reported. Motion compensation requires (1) image registration and (2) pattern orientation correction. Since SIM has line patterns superimposed, it is not easy to register raw frames directly. By introducing two widefield

frames, bookending the three frame sequence, it is possible to compensate for the shift. By compensating for the shift, however, line pair misalignment is introduced. In order to correct for misalignment, by using the high-speed pattern shifting capabilities of the DMD, we reorient the lines parallel to the motion. Shift and misalignment correction then led to real-time mosaicking using SIM endomicroscopy.

Residual line artefacts, both in-focus and defocus plague SIM images. They come from imperfect bundle sampling of modulation. The lines in raw images can be decomposed into a Fourier sine series. Artefacts are caused by the OTF of the bundle dampening and clipping harmonics near the Nyquist limit, meaning the modulation resembles trapezoids instead of lines or sine waves. By introducing random field illumination into endomicroscopy, it is possible to reduce in-focus artefacts with only a slight penalty paid in optical sectioning strength. A pixel shifting approach can be used to reduce artefacts even further. There are additional added benefits to using non-sinusoidal patterns. In SIM, reconstructed image contrast is also limited by sampling effects of a bundle. In Chap. 6 preliminary results are presented showing it is possible to modulate the focal plane with 3x higher spatial frequency modulation and maintain contrast transfer.

Line-scanning confocal endomicroscopy is a high-speed (> 100 frames/s) alternative to point-scan confocal endomicroscopy, but light returning far-from-focus degrades images. Subtraction based approaches exist which suppress this but require multiple frames. Two frames are needed, one with the detector slit and illumination slits aligned and another with the slits lateral offset. The consequence of multi-frame acquisition with a single camera is a halving of the frame

rate and motion artefacts. To eliminate these downsides, in Chap. 7 a high-speed subtraction based approach using a single frame is developed. A modification is reported that allows the rolling shutter of a CMOS camera to act as both the aligned and offset detector slits, and acquire the two images needed for the subtraction process in a single frame. Reduced artefacts are demonstrated, and improved axial response at the half-power and 10% power distances are confirmed. The approach is validated with porcine oesophagus, and high-speed mosaicking at 240 Hz is demonstrated.

Chapter 2

Background

2.1 High resolution microendoscopy

A simple endomicroscope can be constructed with a fibre bundle, LED, filters, an objective, a dichroic beam splitter and camera. An early system was reported by Muldoon et al. in 2007 [47, 48]. This fluorescence system is shown in fig. 2.1. The light emitted from the LED is roughly collimated, passes through an excitation filter, is reflected off a dichroic beam splitter, then projected onto the face of a multi-core ($\sim$ 30,000) fibre bundle, discussed in more detail in Sect. 2.2. Light is then transferred through the bundle, in pixelated form, and excites fluorescence in either endogenous compounds such as collagen or fluorescent stains. The emission is imaged from the proximal face of the bundle onto a camera via the objective, dichroic beam splitter and emission filter, and tube lens. Non-scanning LED-based systems use full-field acquisition, and there is limited inherent optical sectioning.

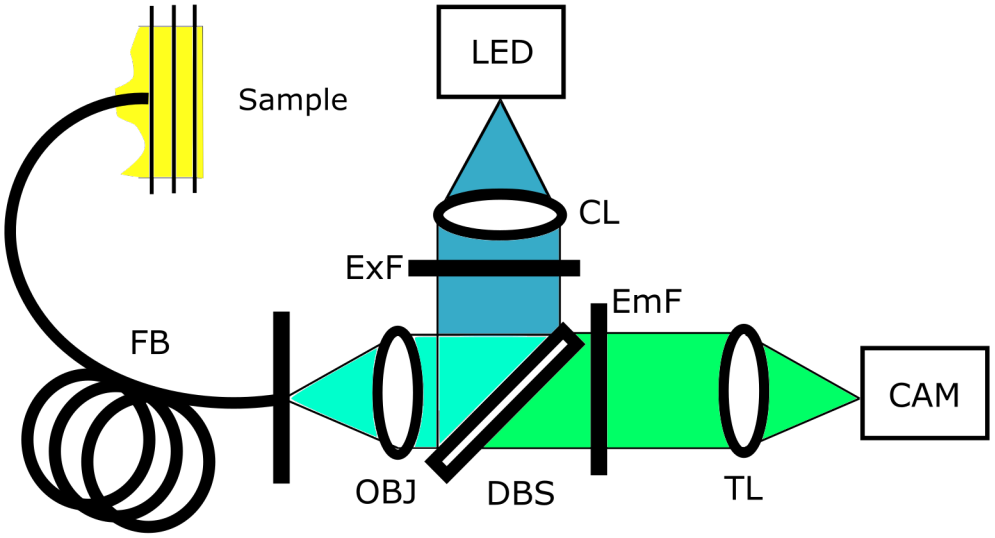


FIGURE 2.1: Scanning free high resolution endomicroscope. Low cost fluorescence system with LED source and multicore (~30,000 core) waveguide. Based on system described in Refs. [1, 47]

Darkfield contrast can also be implemented using the HRME, with fluorescence filters removed, to avoid reflections off the proximal face. The key modification of the LED based system is the introduction of a circular stop in the infinity space region of the excitation arm (Fig. 2.1, this is right after CL) and the introduction of an iris in the infinity collimated portion of the emission arm (before TL in Fig. 2.1).

2.2 Fibre imaging bundles

Multi-core fibre imaging bundles form the core of several endomicroscope designs, including the HRME and the ones presented in this thesis. A quartz fibre bundle is made up of glass rods which are stacked into a hexagonal pattern. The rods are then turned vertically, heated and drawn into a thin bundle, often <1 mm. Fibre bundles consist of a jacket, cladding, and cores. As can be seen in Fig.

2.2 light propagates through the bundle in pixelated form. The Nyquist limit limits the lateral resolution to approximately twice the inter-core spacing (pitch). The resolution also depends on the orientation of the probe. There are a few reasons for this, first, whether or not a feature is aligned with the hexagonal grid of the fibre bundle probe, second, the variation in pitch between the individual fibre cores. Core pitch variation was examined by Chen et al. for a Fujikura FIGH-350S 10,000 core-bundle with a mean pitch of 3.2 μm and was found to have an estimated variation of 7-10% [49]. Cross core coupling, which is wavelength dependent, can be as high as 90% and also diminishes resolution [50]. These factors mean the resolution predicted by the Nyquist limit is often not realised.

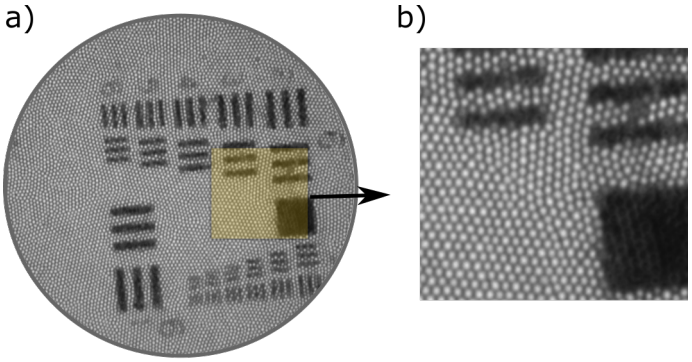


FIGURE 2.2: Image of a bundle face of a multicore fibre optic bundle imaging a USAF target in transmission mode.

The same principles govern light propagation through bundles as that of single fibres. When light propagates through the cladding, it can lead to a deterioration of contrast in the resulting images; this will be seen to be the case experimentally in Sect. 4. Furthermore, the coupling angle of light is manifest as a spatial intensity distribution in each core at the receiving end of a bundle. The modal information encoded in each core presents additional challenges when trying to

remove the cladding structure from raw frames in the spatial frequency domain; this is discussed in depth in Sect. 3.1.

2.2.1 Light propagation through fibre

The theory of fibre optics presented in this subsection is adapted from Snyder et al. [51] as well as Burmel et al. [52]. Fibre optics consist of three primary components: the core, the coating and the cladding. In the geometric theory of light, a ray will propagate through a fibre optic when the criteria for total internal reflection are met. The core must have a higher refractive index than the cladding. The reason for this is shown in Fig 2.3(a) which is a diagram of a ray incident on a single core fibre optic. A large portion of the light is lost to the boundary when $n_2 < n_1$. θ_c is the critical angle, and it is the angle of incidence on the interface at which total internal reflection begins to occur. Total internal reflection is only satisfied when $\theta_i > \theta_c$, this is one reason fibres have a maximum bend radius since each bend effectively changes θ_i. Once θ_i falls below θ_c, there is no longer total internal reflection. Once the critical angle is known, θ_a can also be determined. θ_a is the maximum angle of incident light that the fibre bundle will accept, taking the sine of this gives its numerical aperture (NA). The sine of the coupling angle is the NA when θ_i=θ_c. In air the formula to determine the NA is given by:

$$NA = \sqrt{n_1^2 - n_2^2} \tag{2.1}$$

if n_1 and n_2 are close to each other than $(n_1 + n_2) \approx 2n_1$:

$$NA = \sqrt{(2n_1)(n_1 - n_2)} \approx n_1\sqrt{2(1 - \frac{n_2}{n_1})} \approx n_1\sqrt{2(\xi n)} \tag{2.2}$$

where ξn is $\left(1 - \frac{n_2}{n_1}\right)$.

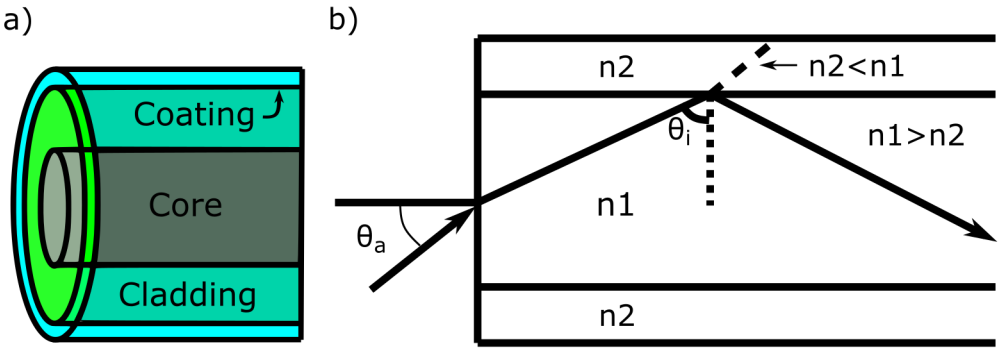


FIGURE 2.3: (a) layers of a single core fibre optic. (b) Refractive index changes at interface layers using geometric optics.

Losses in fibre can come from a few sources including Raleigh scattering at lower wavelengths, at 1.38 um there is a silica OH absorption peak, and beyond 1.6 um the infrared absorption of silica begins to dominate in this region and losses begin to climb sharply. Loss in a typical silica-based fibre is dominated by Raleigh scattering and is on the order of 5 dB per km with a wavelength of 750 nm, higher for shorter wavelengths, however, with fibres only 1-2 m in length these losses are negligible.

Geometrical models only carry so far without needing concepts from a more robust wave theory. One example of this is the evanescent field which penetrates the cladding of a fibre optic. Geometric optics would predict that any penetration of light into the cladding must come from a loss of total internal reflection in the fibre, but this is not lost; it is simply field which permeates the boundary. The wave theory predicts modes of propagation.

The nature of the work in this thesis involves imaging the proximal face of a fibre bundle onto a camera. In each core, there is a spatial distribution of intensity, corresponding to modes. For the work in this thesis, this spatial intensity distribution in each core is averaged over either inherently by camera sampling, or intentionally via cladding suppression routines (Sect. 3.1). A full derivation of modal propagation can be found in Ref. [52]. However, brief discussion is warranted.

Maxwell's equations for a nonconducting medium, and relationships between electric and magnetic flux densities are used to derive the wave equation:

$$\nabla^2 \tilde{E} + n^2(\omega) k_0^2 \tilde{E} = 0 \tag{2.3}$$

where, $\tilde{E}$ is the Fourier transform of the electric field (E), n is the refractive index, ω and k_0 are related to the wavelength of light (λ) by:

$$k_0 = \omega/c = 2\pi/\lambda \tag{2.4}$$

There are a number of requirements to use this wave equation to solve for modal propagation. First, the wave-guide must be nonmagnetic (no magnetic flux density) and without free charges, second, the optical frequency must be much greater than the resonance of the medium, third, the fiber must have low loss (the optical frequency dependent dielectric constant must be real), fourth, the refractive index in the core and cladding must be independent of radius [52]. The wave equation can then be written in cylindrical coordinates, and the method of

separations of variables can be used to obtain three ordinary differential equations [52]. After imposing boundary conditions and constraints the only allowed solutions correspond to the allowed modes [52].

One result of the wave theory leads to a well-established criterion for determining the number of modes allowed in a fibre. The normalised frequency is given by [52]:

$$\nu = \frac{2\pi d_c}{\lambda}\sqrt{n_1^2 - n_2^2} = \frac{2\pi d_c}{\lambda} NA \tag{2.5}$$

In this equation d_c is core diameter. When ν is below 2.44, the fibre is in single-mode operation, when it increases the fibre begins to permit more modes. A table showing the modes, at what frequency they become active, and the total number of modes active is shown in Tab. 2.4, each mode has two polarisations.

Modes	(ν)	Total # of modes
LP_{01}	0-2.4	1
LP_{11}	2.4-3.8	3
LP_{21}	3.8-5.1	5
LP_{02}	3.8-5.1	6
LP_{31}	5.1-5.5	8
LP_{12}	5.5-6.4	10

FIGURE 2.4: Table of modes excited at different ν values.

Different modes travel with different fractions of the field in the core or the cladding. The V number also determines the relative power in the core and the cladding. A V number above 1 is required for at least 50% of the power to travel

in the core. At a V number of 2.4, $\sim 90\%$ of the power is transmitted through the core. The spatial manifestation of higher-order modes is seen in images, techniques to computationally filter them are discussed in Sect. 3.1.

From a manufacturing point of view, it may be tempting to make the cladding as thin as possible to minimise the materials used. However, this does have limits and a cladding with too small of a diameter is more susceptible to loss since the field extends beyond the core and into the cladding. It is, for this reason, there are several reasonably standard single-core multi-mode fibre diameters $\sim 50 - 65$ um, with a typical cladding diameter of 125 um. In the case of fibre bundles sub-optimal spacing between cores leads to light crossing between neighboring fibre cores.

2.2.2 Cross-core coupling

Fibre bundles are subject to the effects of cross-core coupling manifest as a blurring of images [53]. Chen, X. et al. experimentally determined the cross core coupling effects of two different fibres with pitch 3.2 μm and 4.5 μm [49]. At specific wavelengths, it was possible to get as much as 90% cross-core coupling. Specific wavelengths do exhibit coupling minimums, near 0% coupling, however, as little as 1/2 nm shift can lead to a complete reversal in the coupling behaviour. Experimental results also confirm there is as much as a 20% power shift between adjacent cores introduced by the polarisation of the incident light [49, 54]. Polarisation based methods to compensate are challenging to implement in practice since they are often scrambled by the bending of the fibre bundle. Methods to compensate it has been explored [55, 56]. Although the resolution depends on

the pixel pitch, it also depends on cross core coupling. Chen, et al. showed at a higher wavelength (968 nm), the lateral resolution of the fibre with 4.5 μm pitch outperformed the 3.2 μm since it suffered from fewer cross core coupling effects. One proposed solution to reduce the effects is a deconvolution [57].

2.2.3 Reflections off proximal face

The majority of endomicroscopy systems presented in this section are fluorescence-based. One reason for this is image contrast, a second is that back reflections off a fibre bundle overwhelm reflectance signal [58]. Background subtraction does work to attenuate these reflections; however, it leads to a limit dynamic range. Another approach to limit reflections is via refractive index matching using a wedge separated from an angle cleaved fibre with glycerin (n=1.47) [59], however, this is unsuitable for widefield reflectance applications [58]. Another approach involved using illumination outside the bundle; this has been demonstrated in an endocytoscopy applications [60], which is the process of using the absorption and scattering of stains similar to histology, although it does not provide true reflectance contrast [60], this approach is discussed in more depth in Sect. 2.7.3. External illumination has also been demonstrated as a way to obtain phase contrast with oblique black plane illumination [61, 62]. Light injection near the distal tip with light travelling via the cladding has also been proposed [58], the limiting factor is that confocal or structured light methods cannot add optical sectioning since they all require control of the spatial distribution of light on the proximal face.

2.2.4 Distal optics

Fibre bundle endomicroscopy configurations can be lensless or have distal lenses. Distal lenses change the magnification but not the proximal sampling of excitation light. All results presented in Sect. 7 4, 5, 6 are done with a lensless bundle, however, the results could be directly applied to a lensed configuration. The bare proximal fibre face also means that probes can be interchangeable, assuming there is no scanning at the distal end. Single-fibre endomicroscopes require distal scanning with miniaturised scanning elements. Bundle based probes are more amenable to miniaturisation since there are no distal elements required, and those that are optional such as lenses, or focusing mechanism can often be incorporated in a way to maintain a thin form factor.

Distal lens elements can be designed to increase the resolution of a probe at the expense of field of view or focus deeper in a sample. Secondary reasons included correcting aberrations, minimising reflections, co-registering fluorescence for multispectral imaging approaches, or guiding light in exotic ways. The commercial system which uses a bundle as its core with distal magnification elements is the Mauna Kea Cellvizio. The Cellvizio, described in more depth in Sect. 2.4.2 uses proximal scanning across the face of the fibre bundle with a galvanometer for the slow axis combined with a resonant scanner for the fast axis. A summary data table of available Cellvizio probes is presented in Tab. 2.5. While Mauna Kea does not report the umbilical length or diameter, the rigid tip diameters appear to maintain endoscope compatibility across a broad range of probes. The axial and lateral resolution also make them attractive options for confocal fluorescence applications. Plastic achromatic distal lenses have also been reported for

	Lateral resolution (μm)	Axial resolution (μm)	Working distance (μm)	Rigid diameter (mm)
S1500	3.3	15	0	1.5
S300	3.3	15	0	0.3
S650	3.3	15	0	0.65
CerboFlex™ J	3.3	15	0	0.47
Z 1800	3.5	70	100 (488 nm) 170 (660 nm)	1.8
MiniZ	3.5	30	50 (488 nm) 70 (660 nm)	0.94
UltraMiniO	1.4	10	60	2.6
MiniO/30	1.4	10	30	4.2
MiniO/100	1.4	10	100	4.2

FIGURE 2.5: Summary data table for Mauna Kea Cellvizio probes. Based on information provided in Ref. [64].

fluorescence systems [63] capable of 2x magnification.

2.3 Confocal microscopy

The gold standard endomicroscopes commercially available use the confocal principle to obtain optical sectioning. The confocal principle and theory are presented in the following section.

2.3.1 Original patent

Marvin Minksy invented the confocal microscope in 1957. It involves both a scanning point source and scanning pinhole. Light is focused through a pinhole and is imaged via a lens onto a point in the sample. Either through reflection, transmission or excitation of fluorescent molecules, light emanates from that point and is focused back through a pinhole, in a conjugate plane, to a photodetector. Any out-of-focus light is physically blocked, axially displaced light is enlarged

over the pinhole. To reconstruct optically sectioned images, the sample needs to be translated laterally, to reconstruct volumes it must also be translated axially. Minsky's original design is shown in Fig. 2.6. In this design, the sample itself is translated axially and laterally.

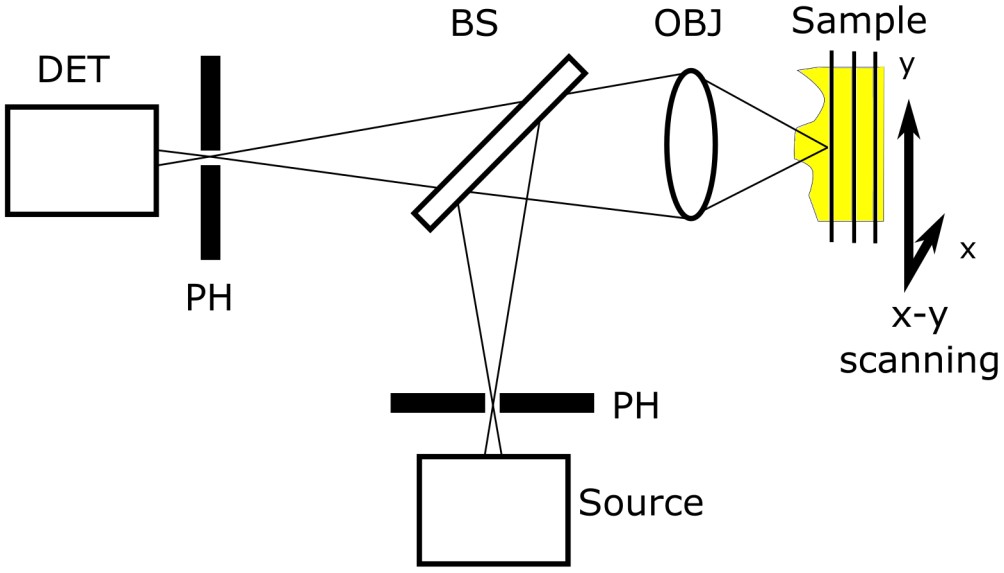


FIGURE 2.6: Marvin Minsky's original confocal microscope. Light guided from an incoherent source to a sample, returning light is imaged onto a pinhole and detector. Based on the system presented in Ref. [13].

Minsky's original design is limited in that the translation stage limits the scanning speed. A modern configuration uses a galvanometer scanner to scan a spot across an image plane. As can be seen in Fig 2.7 [65, 66] a collimated beam is relayed from a laser source, reflected off a dichroic mirror and imaged onto a focused point. Light returning which has been offset laterally is blocked by the pinhole, while light which is an axially different position is attenuated since the intensity is concentrated around the pinhole and only a small fraction return.

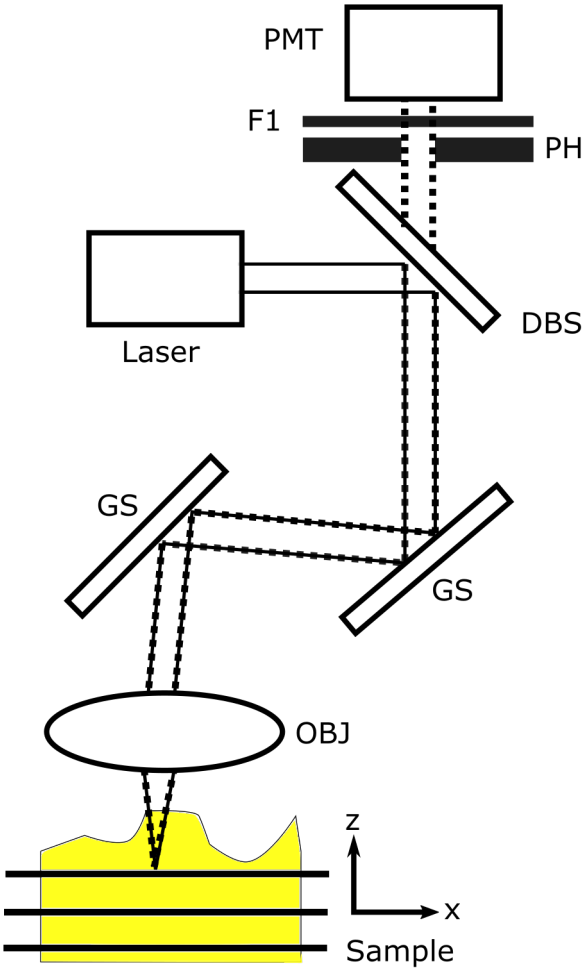


FIGURE 2.7: High-speed galvanometer based confocal microscope. Light is guided from the source to the sample and returns via the same path to the pinhole detector. Based on description in Ref. [65]

2.3.2 Theory

The resolution of a confocal microscope is related to the NA of the objective. With lateral resolution determined by $\alpha\ 1/NA$, and axial by $\alpha\ 1/NA^2$. The behaviour of laterally and axially displaced light can be seen in Fig. 2.8.

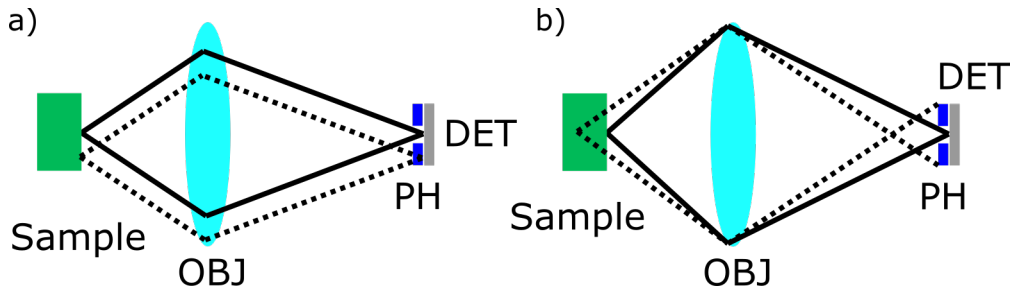


FIGURE 2.8: (a) Laterally displaced rays are blocked from the detector by the pinhole. Axially displaced rays are enlarged over the pinhole.

In the ideal case with an infinitely small pinhole, the formula for lateral resolution is given by:

$$Res_{lateral} = \frac{\lambda}{2\,NA} \tag{2.6}$$

while the formula for axial resolution is given by:

$$Res_{axial} = 0.64\frac{\lambda}{n - \sqrt{n^2 - NA^2}} \tag{2.7}$$

where n is the refractive index of the imaging medium (e.g. air, oil), NA is the numerical aperture of the objective, and λ is the wavelength. There are several complications to the confocal system as presented. First, there is a tradeoff between optical sectioning and signal. The pinhole size determines the quality of the optical sectioning; it also determines the proportion of signal that can return to the detector. The axial illumination cone as well as the excitation of neighbouring in-plane fluorophores contributes to photobleaching.

While the theory predicts an infinitely small pinhole will give the best optical sectioning strength, it will also yield an infinitely small signal. As the pinhole radius fall below the first minimum of the Airy disk, the effect had on optical sectioning diminishes [67]. Axial and lateral response to pinhole radius can be

v_p	Lateral	Axial	Signal % (Fluorescence)
0†	1.2	2.6	0
2†	1.25	2.6	30
4	1.45	4	80
6	1.6	5.5	85
8	1.6	7	90
10	1.6	8.5	95

FIGURE 2.9: Representative lateral, axial resolution (normalised optical coordinates), and relative signal levels as a function of pinhole size ($v_p = kr \sin\alpha$). †below airy disk first minimum (≈ 3.8).

characterised using lateral and axial normalised optical coordinates:

$$u = \frac{2\pi n}{\lambda} x \sin(\alpha), \quad v = \frac{8\pi n}{\lambda} z \sin^2(\alpha/2) \tag{2.8}$$

with u, v being the normalised coordinates for lateral coordinate x and axial coordinate z respectively, λ being wavelength, n being the refractive index and $\sin(\alpha)$ being the numerical aperture. The first Airy minimum occurs when, v_p, the pinhole radius ~ 3.8. When $v_p > 4$ the minimum resolvable feature size is $u = 1.6$. When $1.5 < v_p < 4$ there is a linear relationship between the minimum resolvable feature size and pinhole radius. Below $v_p \sim 1.5$, there is minimal improvement in lateral resolution with further reduction in v_p. The axial resolution as a function of pinhole radius plot has two distinct regions, $v_p < 2.5$ is flat and optimised, and $v_p > 2.5$ the HWHM of axial resolution linearly increases with each radial increase in pinhole size.

To give an indication of the interplay between pinhole radius, lateral, and axial resolution and relative signal a number of lateral resolution, axial resolution values, and signal values are tabulated (Tab. 2.9) as a function of normalized pinhole radius. The data is compiled, from Refs. [67–69].

2.3.3 Fibresation of benchtop systems

Contemporary confocal endomicroscopes result from successive modifications of Minsky's original design. One such of these was transitioning components from free space optics into fibre optics. While, as can be seen in Tab. 2.9, it can be useful to have a wheel of pinholes to vary the signal level and optical sectioning strength there are many circumstances where it is desirable to separate bulky components from an optical system, such as to make the system more compact. The approach that enabled this is known as the fibre optic confocal scanning microscope(Fig. 2.10) [68]. The pinholes on the two arms have been replaced with a source fibre, coupled into a laser source, and a detection fibre . The constraint this does impose is the pinhole size cannot be further customised to balance the signal, and optical sectioning strength since the diameter of the fibre sets them.

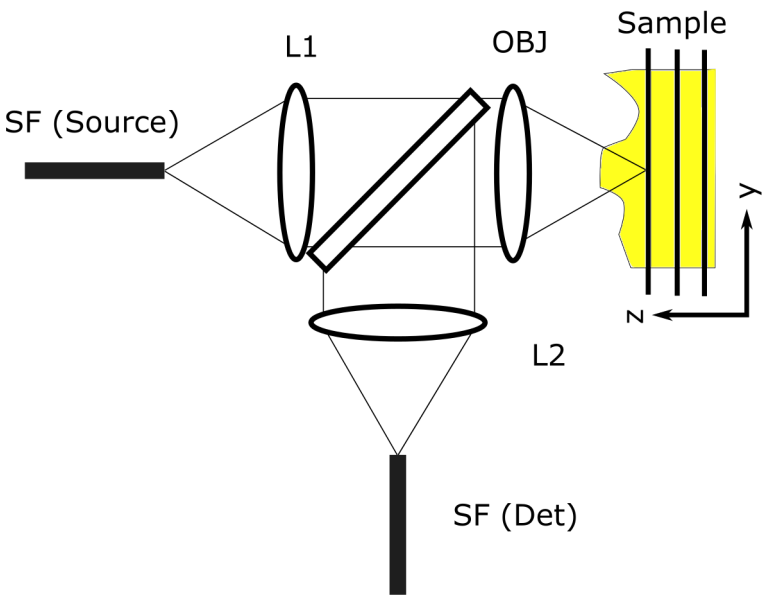


FIGURE 2.10: Modification to benchtop system with detection and source pinholes in fibre. Based on description in Ref. [68]

Dabbs, T. et al. made a further simplification, et al. (Fig. 2.11) The beam splitter was coupled into fibre [70]. This idea was virtually identical to one patented

by Harris, A. et al. [71]. Dabbs, T. et al. paper was published on 1 June 1992, while Harris' patent was awarded eight days later. It is useful to note from the figure that scanning is done by translating the fibre on the distal end not proximally. The transition from this configuration to a distal scanning endomicroscope appears quite logical, requiring only the miniaturisation of the lenses, and a mechanism of distal scanning.

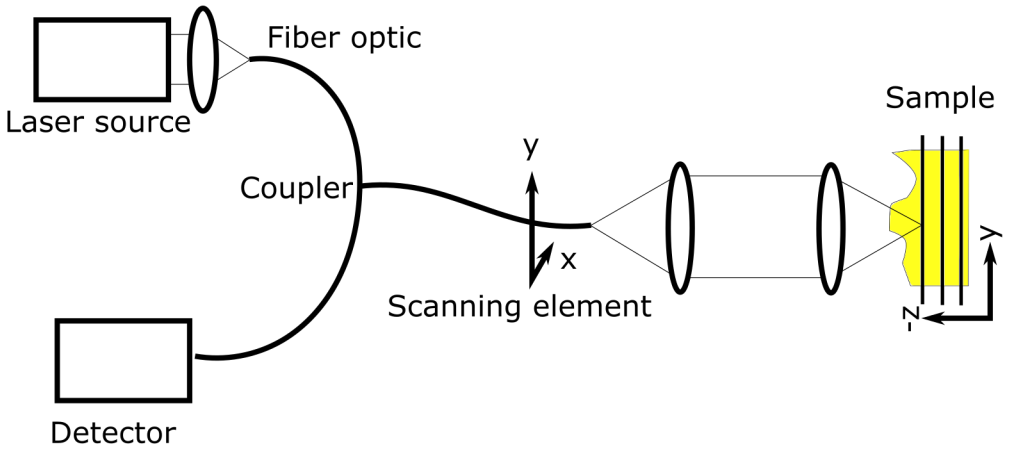


FIGURE 2.11: Distal scanning elements of a fibre based benchtop confocal microscope (Distal elements not miniaturised). Detection and source pinhole sizes set by fibre diameter. Detector and source not shown. Based on description in Ref. [71].

2.4 Confocal Endomicroscopy

2.4.1 CLE with a single core fibre

Endomicroscopes can come in several flavours. They can be built into the end of an endoscope or operate as separate devices which may use the working channel of an endoscope or scan in standalone operation. To maintain endoscope compatibility, it is often desirable to have a distal tip which has a rigid length and diameter that lets it easily traverse the working channel.

In 1991 Giniunas et al. reported a system with miniaturised distally mounted lenses, and the laser alone coupled via a modulator into the single-core fibre [72]. Reflections were then fed back into the laser cavity where a photodiode collected light. Processing electronics then detected modulations in the signal. The entire system was then encased in a box-shaped structure with overall dimensions 4x5x60 mm shown in Fig 2.12. The box consisted of four PZT plates, each two acting as pairs and served to move the entire objective. With the perfect vision of hindsight, this approach has a few apparent limitations, mainly the shifting process is moving heavier components than a fibre bundle, and this would require stabilisation for the PZT elements. The box-shaped structure and size prohibit endoscope compatibility.

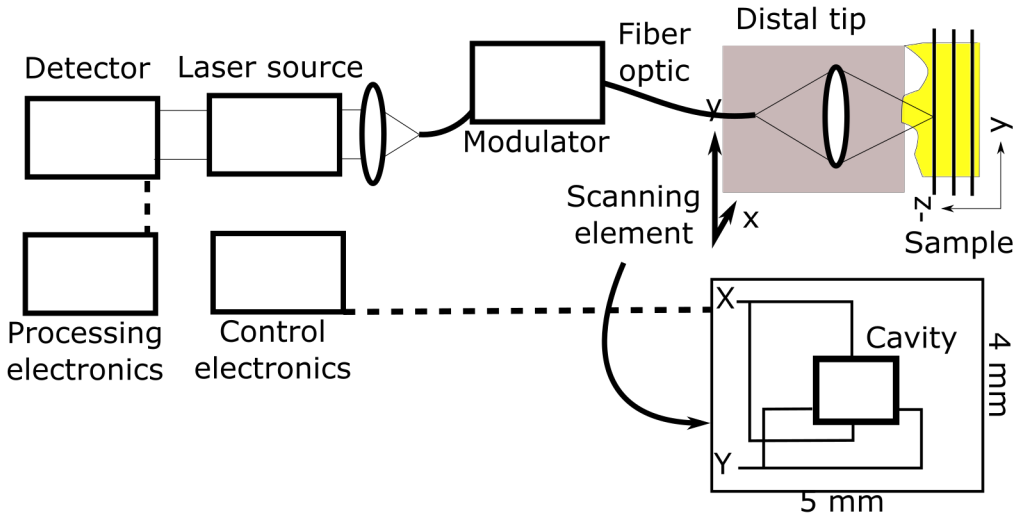


FIGURE 2.12: Early piezoelectric (PZT) based distal scanning single fibre endomicroscope. Back port of laser source used to feed light into detector. Square cavity for probe, distal PZT elements control scanning. Based on system described in Ref. [72].

The next modification to Minsky's original patent was the miniaturisation of the distal optics completed by Delaney et al. in 1994 [73]. Delaney also introduced a focusing system which led to a confocal endomicroscope using a single

fibre [73]. Problematically, distal elements such as focus control or scanning elements as well as lenses increase the amount of space required in the distal head. Endomicroscopes can be used in standalone mode as well as in an endoscope. One such standalone case is when a pathologist examines external tissue such as skin or recently excised tissue interoperatively, such as during a minimal excision (Mohs) procedure. In this configuration, there are no significant limitations to endoscopic head size. For those applications a system based on distal scanning has been commercialised such as the Optiscan FIVE2 [74]. A schematic of the device is shown in Fig 2.13. The lateral resolution, in this case, is limited by the precision of the scanning element in the distal head.

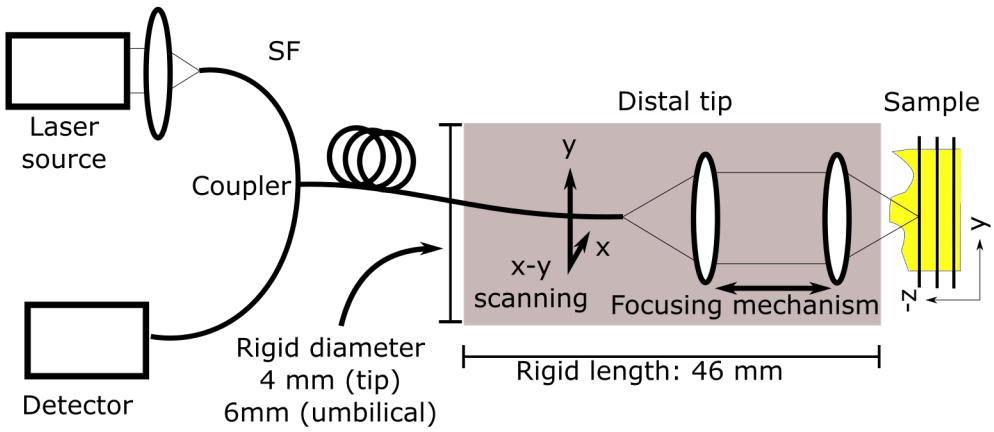


FIGURE 2.13: Single fibre based endomicroscope with distal scanning elements and focusing. Scanning occurs distally in rigid tip, umbilical contains single-core fibre and electrical wires for distal scanning elements and focusing mechanism. Based on design presented by Ref. [73]. Probe diameter and umbilical length prohibit operation in typical 3.2 mm endoscope working channel.

2.4.2 CLE with a fibre bundle

It is possible to reduce the constraints imposed by the sizeable distal head by moving scanning to the proximal end, out of the patient's body. A proximal scanning system (Fig. 2.14) was first reported by Gmitro et al. in 1993 using a

multi-core fibre optic bundle [75]. A fibre bundle is mounted at the object plane of a galvanometer based benchtop confocal microscope, and the flying-spot is raster scanned across the proximal face (Fig 2.7). An advantage of this approach is that the relocation of the scanning element to the proximal end allows a more miniaturised distal tip. The design first reported included a lensed tip, although this is not a requirement. The fibre bundle distally can be used in contact mode, where it is placed in direct contact with the tissue, or it can be used with a lens configuration.

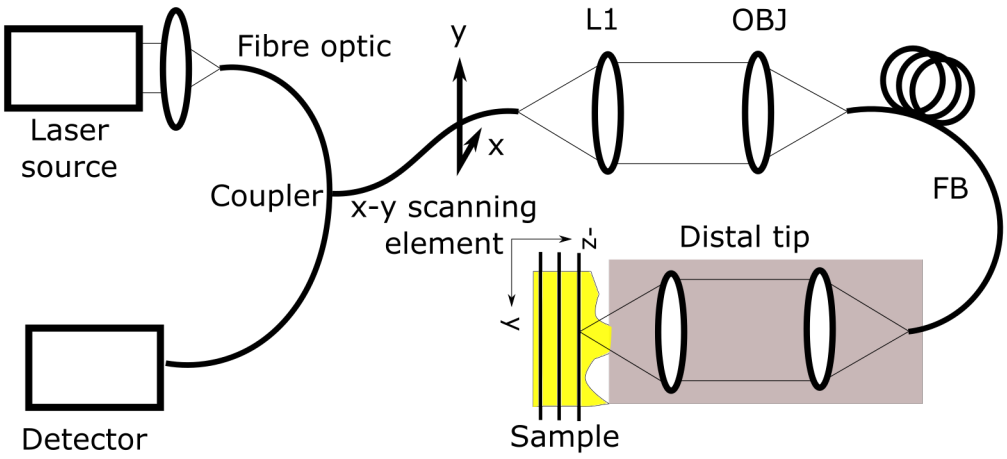


FIGURE 2.14: Fibre bundle based endomicroscope. A single fibre is imaged onto the proximal face of a multi-core bundle. Slow scanning is done with a galvanometer scanner, the fast axis with a resonant scanner. Based on design presented by Ref. [75]

Lensed configurations can include adjustable tradeoffs between the field of view and resolution. Mauna Kea Technologies has commercialised a fibre bundle system as the Cellvizio. The lateral resolution, set by the inter-core spacing of the pixels in the fibre bundle, can be improved at the expense of the field of view through the selection of distal lens elements.

2.4.3 CLE with FB and a Nipkow disk

Single point scanning systems are often slow. A proposed way (1997) to improve speed is with a Nipkow disk acting as a scanning pinhole, this was the first real-time demonstration of confocal fibre bundle endomicroscopy [59]. A Nipkow disk is shown in Fig 2.15 (b). It is a metal disk with holes placed at different radii and angular positions. The region of the disk enclosed in black would be in a region conjugate to both the source detectors, and bundle surface, shown in Fig 2.15 (a). The disk provides the point source as well as point detection. The Nipkow disk is rotated, and snapshot images are acquired. Based on the system described, it likely resembles the configuration displayed in Fig 2.15. The configuration is a white light confocal system which is reflectance-based. The authors chose a white light source to avoid the levels of speckle, deemed unacceptable, that were manifest in a laser-based non-fluorescence configuration.

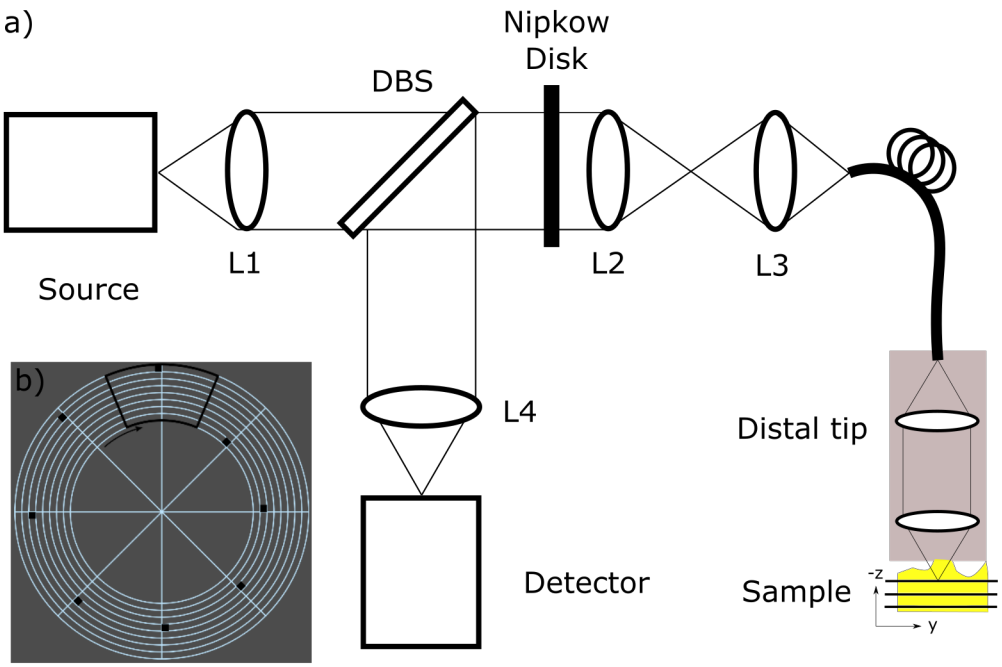


FIGURE 2.15: (a) Schematic of a fibre bundle based endomicroscope with a (b) Nipkow disk. A fibre bundle is mounted at the sample plane of a Nipkow disk confocal microscope. Based on the system described in Ref. [59].

2.4.4 CLE with a fibre bundle and DMD

In 2000, Lane et al. reported a non-fluorescence approach to achieve confocal sectioning using a DMD placed conjugate to a fibre bundle to confine the illumination to a single core [76]. When only a single core was illuminated, and emission light returned to adjacent cores, it was known that this light must have been scattered and was filtered as out of focus. A diagram showing a conjugate plane to both the DMD and fibre bundle is shown in Fig. 2.16. The DMD was registered, so that 6 DMD pixels were used to illuminate a single fibre core. Previously light would illuminate the whole bundle, including the cladding; this contributed to reflections and a low SNR.

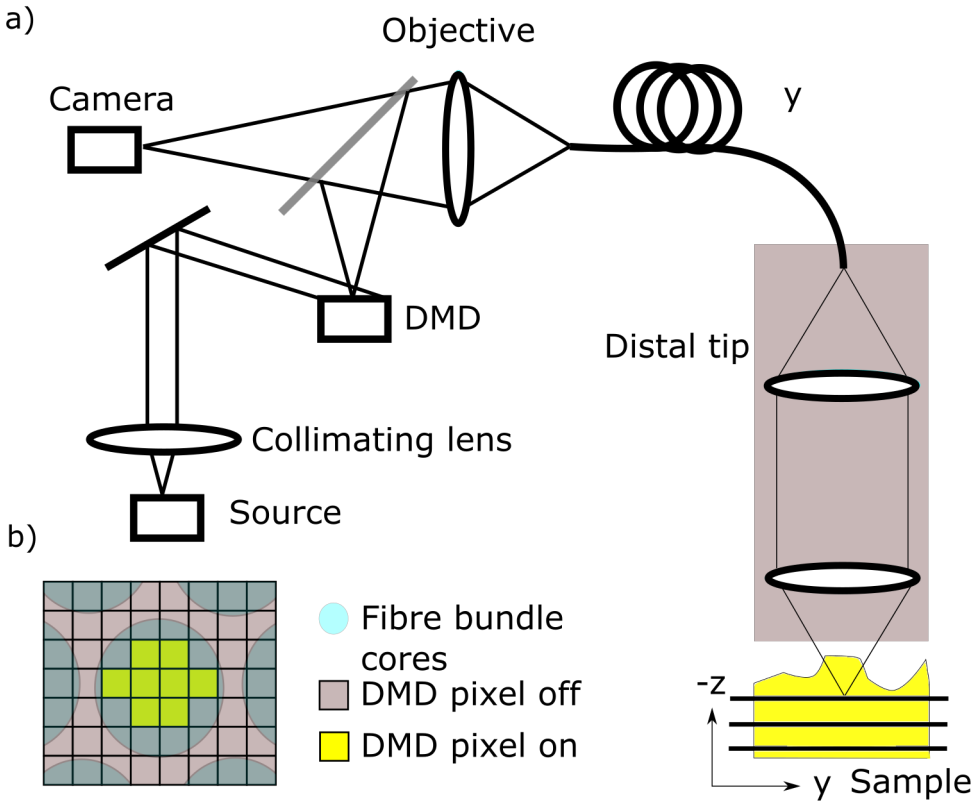


FIGURE 2.16: (a) Schematic of a fibre bundle based endomicroscope with DMD illumination of each core. DMD is conjugate to proximal face of fibre bundle. 6 DMD cores enclosed completely in the area of each fibre bundle core. (b) Illustration of single core illumination via DMD pixel to bundle core registration. Based on system and illumination approach developed in Ref. [76].

Lane et al. reported when an approach where the cores of a fibre bundle were registered to multiple DMD pixels [76]. By only illuminating a single core of a bundle with around 6 DMD pixels there was 30% efficiency in transmission, an improvement over 18% when core-specific coupling is not employed. The diagram as presented has no inherent optical sectioning but behaves instead like the microscope with only the point source. The lateral resolution is determined by the separation between the cores, and axial from the NA of each fibre core however light returning to neighboring cores limits image quality. To further improve axial resolution, Lane et al. used camera masking. Using the known pixel mapping from the DMD to the camera they discard light returning from non-illuminated cores. The DMD has a 15 us delay from the time a pattern change instruction is received by the DMD to the time the mirror is active. The delay imposes a small, but non-negligible, time constraint. For single-core illumination DMD limited full bundle acquisition would take 0.45 seconds per frame or 2 frames/s. The required camera frame capture rate would also need to be 66 kHz making it impractical. To recover speed multiple cores can be illuminated simultaneously the cost is out-of-focus light can return through multiple detection cores. Wang, Q et al. (2016) reported one such benchtop configuration [77] of multi-spot illumination. It is of course much faster; however, it remains a multi-frame approach making it subject to motion artefacts.

2.4.5 Line scan confocal

Line-scan confocal, described in more detail in Chap. 4, involves the scanning of a detection and illumination slit across the proximal face of a bundle. Unlike point-scan confocal, line scan involves scanning in only one dimension making it faster. An endomicroscopy implementation was first reported by Sabharwal (1999) using a galvanometer scanner and physical detection slit. It was later reported uses a linear CCD for high speed ($>$ 100 Hz) acquisition [78]. A subsequent modification used the rolling shutter of a camera as an electronically variable detection slit [79]. A DMD approach has also been used for line scan endomicroscopy. The DMD implementation used rows of a DMD as electronically variable illumination gates [80]. The DMD has limited on-board storage limiting the minimum line width, and its fixed pattern change rate was less than the sweep speed of the camera. Parallel illumination is reported to overcome the storage limitation, however, the line width is limited by the pattern change speed [81].

A problem with line scan confocal is that far-from-focus light is not filtered out and degrades images. Consider a line scan system with a small lateral offset between detection and illumination slits, this leads to the rejection of a substantial amount of in-focus light, but only a small amount of far-from-focus light. A subtraction based approach can be used to reduce far-from-focus light. The subtraction approach works using two images, one acquired with aligned slits, and one with offset, and subtracting the image acquired with offset slits from the one with aligned [82]. This was implemented as a multi-frame approach in a galvanometer system [79], and a different approach using two cameras [83]. The problem with the multi-frame approach is it leads to a halving of the frame

rate, and motion artefacts when the probe shifts between acquisitions. This leads to images that don't faithfully reseamble the ground truth. It is often useful to stitch adjacently acquired images into a mosaic, a lower frame rate caused by a multiframe approach limits the maximum speed at which a probe can travel.

2.4.6 Side viewing CLE

Forward viewing probes all suffer from the same limitation; it is difficult to image the sidewalls of cylindrical organs. As a model consider a clinician attempting to image the sides of a cylinder (Fig. 2.17). The fibre bundle tip must bend which introduces several practical problems. Bending can cause stress on the fibre, a loss of total internal reflection or the organ size may be narrower than the distal tip of the fibre allows, components are also needed to bend the fibre.

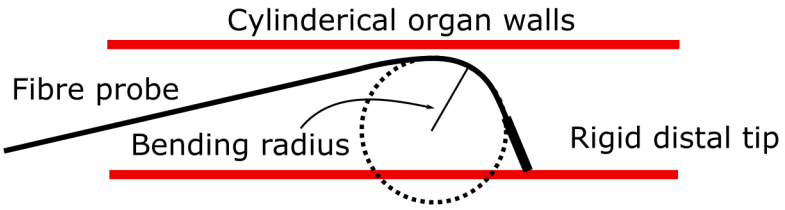


FIGURE 2.17: Forward facing fibre bundle imaging cylindrical organs. Maximum bending radius determined by sharpest bend in fibre.

A rigid type side-viewing confocal endomicroscope such as the one demonstrated by Kim (2010) was developed, which partially addressed the forward viewing limitation [84]. The rigid endomicroscope was used to image rat colon since the navigable pathways were much smaller than humans. The authors present a scanning approach where a rigid endoscope, on a rotating platform is immediately mounted at the focus of an objective. The beam is scanned in 2D

across the lateral surface coordinates and then rotated. This was presented in Fig. 1.5 in Sect. 1.2.

The rigid side-viewing configuration, as reported, is not very useful for endomicroscopy for the human GI tract since the rigid length extends the length of the endoscope. A side-viewing configuration based on the point scan system was then implemented endoscopically [85]. The configuration can be seen in Fig. 1.5, it is very similar to the lensed fibre bundle presented above. Scanning is done proximally, and a mirror and objective are introduced to enable side-viewing. The system was capable of 5 frames/s imaging, a working distance of 100 μm and a FOV of 300 x 400 μm^2. One can imagine the probe needs to maintain enough flexibility to bend easily through an endoscopic channel, however enough rigidity to rotate from the distal end. The characteristics of the probe umbilical under rotation, compression and shear mean the probe must behave be more like wire than yarn.

2.5 Structured Illumination

Structured illumination microscopy (SIM) is a technique which can provide optical sectioning at the expense of multiple frames. The technique and theory are comprehensively introduced in Sect. 4. In broad terms, for comparison to the other variants of endomicroscopy, it is a multi-frame process which delivers optical sectioning by modulating the focal plane with high spatial frequency lines. The lines are phase shifted for three total images and an optically sectioned image is reconstructed. SIM is a subtraction-based technique making it inherently

noisier than confocal.

SIM was first reported for endomicroscopy using a laser and the illumination patterns were generated by reflection off a spatial light modulator (SLM). The SLM is conjugate to the fiber bundle face [43]. SIM background rejection depends on the line pair width. Line width determines two factors: optical sectioning strength and modulation depth. Effectively, greater optical sectioning strength leads to a lower modulation depth, which increases the noise. In 2015 Keahey, et al. reported a SIM system which used lower cost LED illumination and motorised gratings to generate the modulation [42]. Chap. 4 will report and characterise a DMD based structured illumination endomicroscope which generates reconstructed images at 40 frames/s with on-the-fly tunable modulation to balance signal and optical sectioning strength.

As Keahey demonstrated, a benefit of SIM is that lower cost LED illumination can be used and scanning is eliminated. SIM is a multi-frame technique and as a consequence it is sensitive to motion artefacts when the probe has shifted between acquisitions. For this reason, mosaics have not been demonstrated in work before that presented in this thesis. Chap. 5 will report a motion compensation approach to generate mosaics.

Since the spatial resolution of a fibre bundle system is fixed by the inter-core spacing of the bundle, gradient index lenses have been reported for SIM endomicroscopes. In fluorescence applications, however, there is a Stokes shift present, when using distal lens elements which are not achromatic, chromatic aberrations leads to the excitation and emission planes not being co-localised, and decreases collection efficiency. In 2013 Kyrish, M et al. reported a custom-built compound

lens for a SIM endomicroscope which corrected for the Stokes shift, co-localising the planes and improving collection efficiency [63]. The outer diameter of the lens system is around 2.1 mm, which was matched to the outer diameter of a 14 gauge core biopsy needle.

2.6 HiLo

A comprehensive treatment of HiLo is detailed in Sect. 5.5. In comparative terms, HiLo microscopy is a multi-frame technique, developed by Ford, T. et al. in 2012 to accomplish optical sectioning [86]. At the acquisition level, it requires acquiring two images, one which has modulation superimposed, and another which is uniformly illuminated. Through post-processing, the images are separately filtered to detect local contrast variations in regions where modulation, via the grid patterns, appears.

As HiLo is a multi-frame technique Ford reports a similar problem as SIM with motion artefacts. Interframe motion contributes to a decrease in spatial resolution, and since this is a 2-frame technique, the authors use a dual shutter camera to acquire two images in close succession [86]. The dual shutter reduces the distance the probe has moved between subsequent images. The delay they report is 5 μs between images. Contrast this to the 8333 μs maximum exposure of a 120 Hz camera, and it is clear why this would reduce artefacts. One apparent drawback of this approach is the need to switch the source of illumination from uniform to the non-uniform grid patterns in 5 us. The authors used a galvanometer to shift a beam of laser light between a mirror and a transmission Ronchi ruling. Lasers

and galvanometers add complexity to an optical system, and Ronchi rulings do not have adjustable spatial frequencies. Additionally, it is useful to post-process in real-time and high and low pass filtering as well as subtraction and division operations can increase processing time.

2.7 Selected systems

Several recent selected systems are presented to give a sampling of the state of the field. Including fibre bundle shifting with a piezo element to improve resolution, aperture confinement to improve the depth-of-field, endocytoscopy, which works on a similar principle to histology with white light image of staining agents, and spectrally encoded confocal which encodes spatial information by wavelength. Spectral encoding led to improved resolution in fluorescence fibre bundle endomicroscopy which is also presented.

2.7.1 Fibre bundle endomicroscopy with a distal piezo element

Fibre bundle shifting is developed to overcome the ICS limited lateral resolution [87]. By taking low-resolution images, acquried by shifting a probe less than an integer pixel raw images can be embedded into into a finer mesh and interpolated. This approach is well studied in the machine learning community [88]. Since multiple camera pixels are used for each fiber core (~ 7 in Chap. 7) The sampling process of the fibre bundle onto the camera is the process of embedding a low-resolution image into a higher precision mesh. Furthermore, the core

location routine outlined in the Sect. 3.1.4 (Delaunay Triangulation) similarly, involves locating the cores to sub-pixel accuracy.

The sampling process of the object by the bundle limits the lateral resolution. A fibre bundle is a pseudo hexagonal array of cores, and during fabrication, glass rods stacked in a hexagonal pattern, and when heated are slightly shifted and are no longer perfectly periodic. It is mostly locally periodic, if one zooms into a region of only a small number of cores the typical hexagonal structure appears. This structure is then sampled onto a camera, often with a few camera pixels per fibre bundle core. The pixelated structure is approximated in Fig 2.18. This approximation is not strictly correct, due to the modal nature of the bundle, however, after cladding removal it becomes a very close approximation when used with coherent light. Multimode fibres have different spatial distributions depending on the phase and intensity variations of the incident light. Except for work done to exploit this to either improve resolution by numerically confining the aperture [89] or extract 3d information [90] this is mostly averaged over.

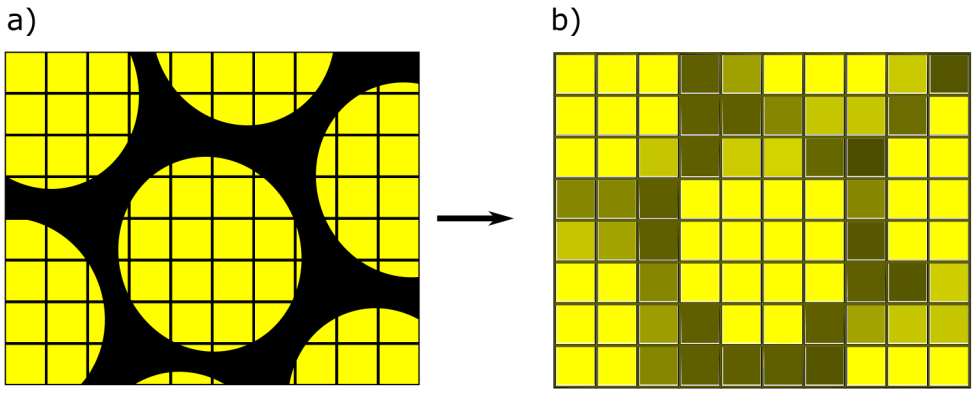


FIGURE 2.18: Fibre bundle core sampling. (a) proximal face of bundle, (b) bundle when sampled by camera.

Once the pixelated structure is approximated, the next step is to remove the residual cladding. Cladding removal is discussed in more depth in Sect. 3.1.

In brief, a few ways to handle this are Gaussian smoothing, a median filter, or interpolation. This leads to a blending of the cladding pixels with the adjacent core pixels. The net result is the same: the low lateral resolution information transmitted through a fibre bundle is upsampled into a finer mesh. This in itself does not improve resolution as high-frequency components were not transmitted via the bundle in the first place [88]. To get resolution improvement images need to be shifted by non-integer multiples of the inter-core spacing, this accomplishes this organically since each inter-core distance is often separated by 3 or more pixels.

For image acquisition a multi-core fibre bundle is placed at the object plane of a line scanning system. A piezo element is built-in to the distal tip of the fibre. Fig. 2.19 shows a diagram of the system with the distal tip elements. The fibre bundle is fixed to a piezoelectric element (PZT), the light is then incident on a grin lens and focused on a sample. The piezo element is then shifted collecting various sub-core shift positions of the piezo device. In post-processing, the acquired images are shifted, and interpolated, leading to a resolution improvement of 2-fold.

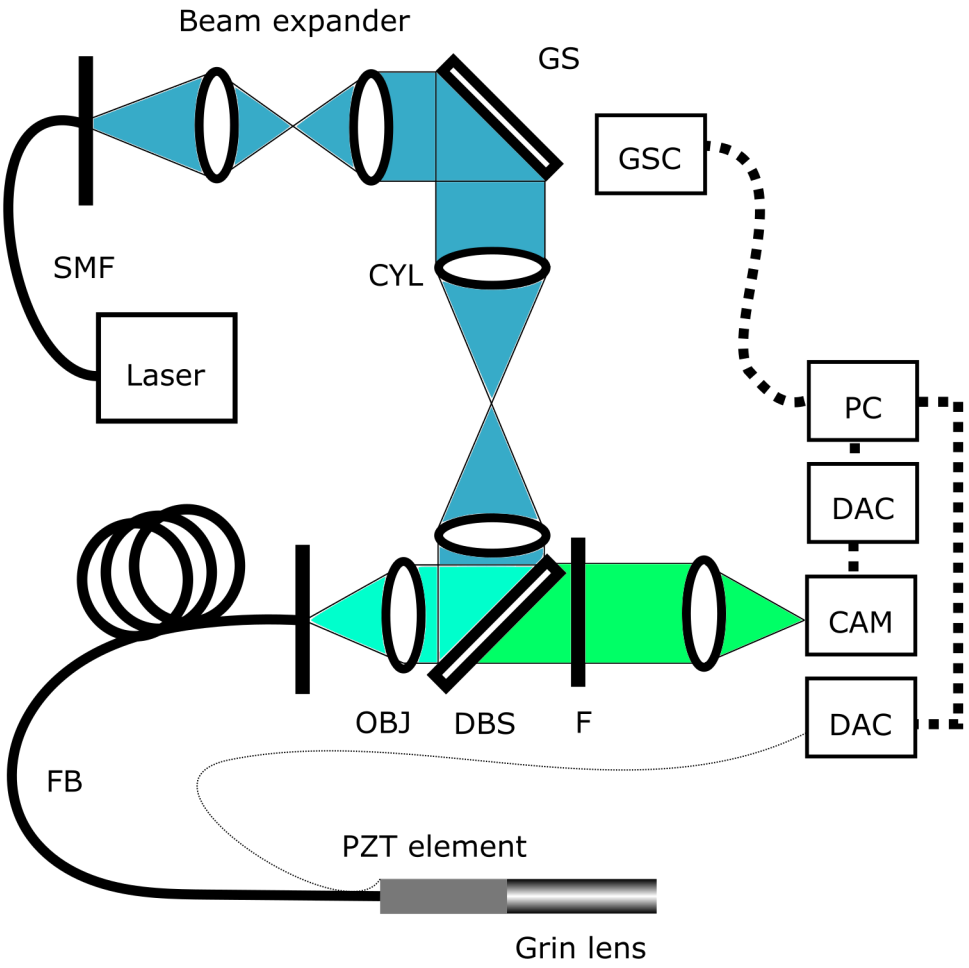


FIGURE 2.19: Line scan confocal fibre bundle endomicroscope with distal piezo element for shifting fibre bundle. The distal piezo element shifts cores by sub-pixel amounts to build up a point map of intensity values. The core system is a line scanning system described in Chap. 7. Based on system described in Ref. [87].

2.7.2 Aperture confinement endomicroscopy

Fundamentally aperture confinement endomicroscopy is the descrambling of spatial information which encodes the coupling angle of light rays into a core to improve image quality. Each core on the fibre bundle no only captures intensity, but encodes directional information in the core as intensity. Typically, intensity is simply averaged over the entire core. Averaging is often done, intentionally

or not, as part of the cladding removal process. This removal of information is true of all the approaches discussed in Sect. 3.1 including Gaussian convolution, median filter, nearest neighbour interactions, Delaunay triangulation, and local area comparison.

Fig 2.20 shows how light incident on a single core from different angles, θ, are seen on the distal end of a fibre bundle. The orientation of the lobes does depend on the value θ, and with perfect cylinder cores, and many camera pixels sampling each core, it is possible to resolve, with 2 fold degeneracy the angle of the incident ray. What is also clear is that a paraxial ray has intensity more centrally located in a core. Meaning light which emanates from greater angles leads to greater intensity at the edges of a core. Furthermore, the coupling efficiency of the light drops as the angle becoming increasingly higher. Inefficient coupling at greater angles means the bundle has a limited depth of field. The authors further exploit this property to artificially crop out the outer edges of the cores, since off-axis light when imaged through a bundle only degrades image quality, and off-axis light preferentially illuminates the edges of the bundle. The authors further employed a subtraction technique using two images, one with the central aperture confined, and one with light collected from the whole core. The images were then subtracted to generate improved depth of field. The approach used a lensless fibre optic bundle and reported a 78% greater depth of field at 50 μm.

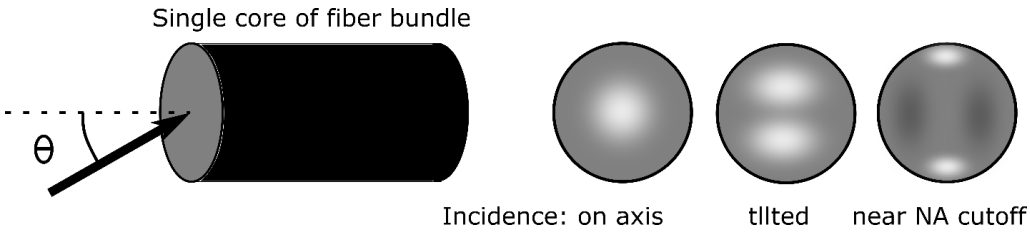


FIGURE 2.20: Example of excitation of higher order modes of a multimode single core fibre. The angle of a light ray incident on a core is shown as $theta$, distal patterns then show spatial encoding at a few $theta$ values.

2.7.3 Endocytoscopy

Endocytoscopy is endomicroscopy using, as a contrast mechanism, the absorption and scattering of white light by coloured stains in a way similar to conventional histology [91, 92]. It was first described in 2006 by T. Ohigashi et al. [93], and patented by Olympus in 2007 [91]. Ohigashi reported the use of a 3.2 mm flexible endomicroscope in the 3.2 mm working channel of a rigid cytoscope. In 2013 Hughes et al. reported a multimodal system where images from a white light endocytic system were combined with a confocal fluorescence imaging system [60]. The confocal fluorescence endomicroscope components were similar to the ones shown in Fig. 2.16, first reported by Lane et al. [76], with the key difference being the use of free space optics.

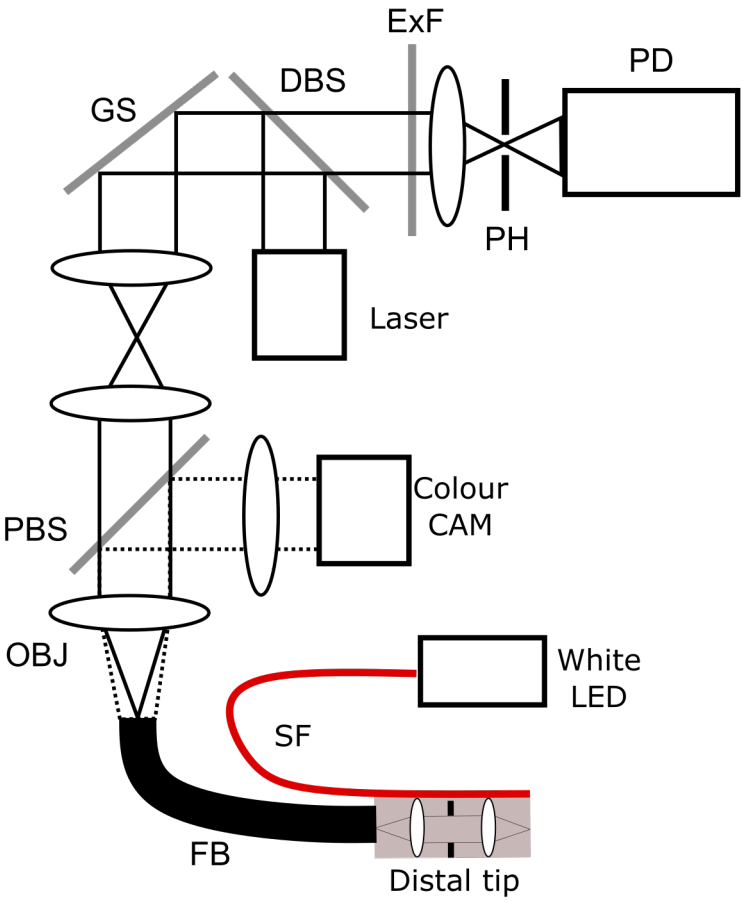


FIGURE 2.21: Endocytoscopy endomicroscope with distal lens elements. A fluroescence configuration, similar to that described by Lane, et al [76] is coupled with a reflectance configuration at the polarising beam splitter. Two alternating images are captured, one reflectance, one fluorescence. Based on system described in Ref. [94].

A diagram is shown in Fig. 2.21, the fluorescence system is separated from the reflectance at the point of the polarising beam splitter, the frames are acquired sequentially confocal then reflectance. The reflectance system involves relaying light returning from the tissue, side-illuminated by the single fibre coupled into the white LED. The light is transmitted through the fibre and incident on a polarising beam splitter. The authors mention that the PBS was used, although light inefficient, instead of a dichroic beam splitter since a dichroic beam splitter would have meant the green light for reflectance imaging would not have returned to

the camera. The remaining operation of the confocal is as described in prior reports [75].

2.7.4 Spectrally encoded confocal with a single fibre

In typical endomicrsocopy spatial encoding of light is accomplished in different ways, depending on the configuration. In the case of a single fibre, the fibre's distal position and the properties of distal lens elements localise the lateral as well as axial source of the light. With a multi-core bundle, the lateral position is determined by the location of the cores and axial by distal focusing elements. Spectrally encoded confocal is a way of storing information in the spectrum of light as opposed to spatially.

An alternative to spatial encoding is spectral encoding. A spectrally encoded confocal technique was developed by Tearney, G et al. in 1998. Spectrally encoded confocal uses a light source with a narrow bandwidth which is collimated from a single fibre, incident on a diffraction grating and the spread out in one dimension over a sample [95]. Each wavelength corresponds to a single lateral position. 1D scanning is required to reconstruct a 2D image. The confocal properties can be seen in the Fig 2.22, only light emanating from focus is coupled back into the fibre. Due to the reversability of diffraction gratings, the decoding portion is simple as well. The only difference from the encoding process is that the sample is replaced with a linear CCD. The resolution is given by the spectral dispersion. Spectral encoding was first reported endoscopically, in a side-viewing configuration using a single fibre in 2007 by Yelin [96]. Scanning was achieved with the introduction of a rotating MEMS scanner after the diffraction grating.

As presented in 2008, this technique is not compatible with fluorescence since the excited fluorophores will necessarily return light at at different wavelength being deflected away from the fibre.

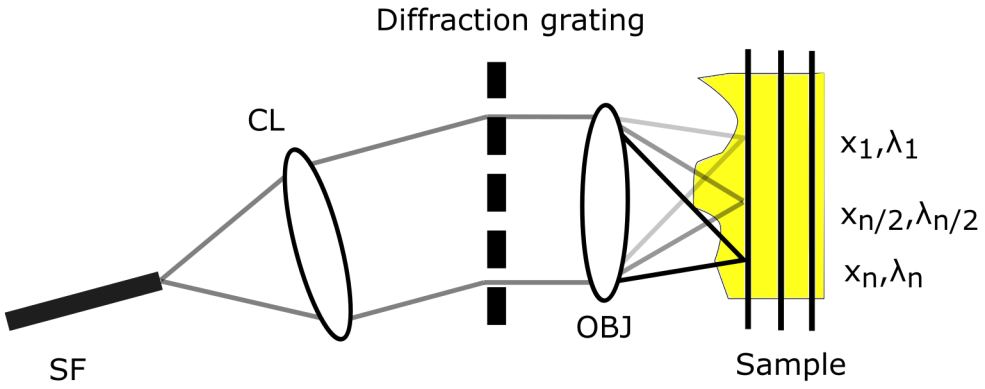


FIGURE 2.22: Principle of spectrally encoded confocal with a single core fibre.

Two methods have been proposed to allow for fluorescence compatibility. The first is the use of a high-speed tuneable swept-source [97] for a benchtop configuration. No endomicroscopy application using a single fibre is known. The second method of fluorescence for benchtop systems was reported in 2005 by Motz et al. using a spectrum dependent modulation and detection [98]. A frequency modulation approach for fluorescence has not been implemented in an endomicroscope.

2.7.5 Spectrally encoded fibre bundle enodmicroscopy

One application of fluorescence spectrally encoded endomicroscopy is in overcoming the lateral resolution limit of a fibre bundle. Light can be spectrally separated at the distal end of each core of a bundle through the use of a wedge prism. The entire bundle is illuminated with individual wavelengths, and snapshot images are recorded [99]. The lateral offset from the centre of the core is known

from the gradient of the prism. The number of snapshots increases the number of pixels. Fig. 2.23 illustrates the principle of this approach. In the same way as fibre bundle shifting, this leads to shifts in the fiber at subpixel positions, which can be interpolated to improve resolution. Spectral encoding was employed with fluorescence. Although fluorescence collapses the spectral encoding, the excitation centres are known from the core position and wavelength.

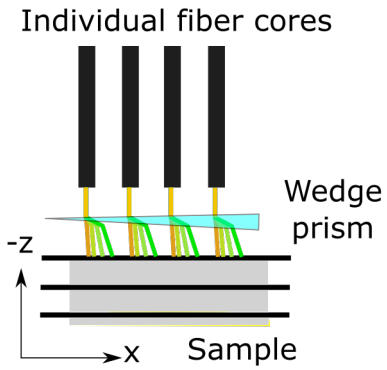


FIGURE 2.23: Light lateral seperation by wavelength using spectral encoding with a fibre bundle and wedge prism.

2.7.6 Two photon fluorescence

In 2019 Kim, Y et al. demonstrated two-photon endomicroscopy as an emerging modality for depth imaging [100]. Conventional endomicroscopy offers limited penetration depth, for applications that require this, endoscopic optical coherence tomography can be used [101]. OCT resolves light reflected at interfaces and is capable of generating volumes [102]. Only the transmission and reflectance properties of the interface are used. In contrast, two-photon fluorescence is linked to specific endogenous compounds or staining agents. Two-photon also offers advantages over confocal techniques due to confocals limited depth penetration and

relatively high photobleaching [103, 104]. Photobleaching is improved since fluorescence excitation only occurs in a region where there is a high enough photon flux, contrallable with choice in a laser source and the NA of an objective. The penetration can be improved using near-infrared light which penetrates deeper in the sample [105].

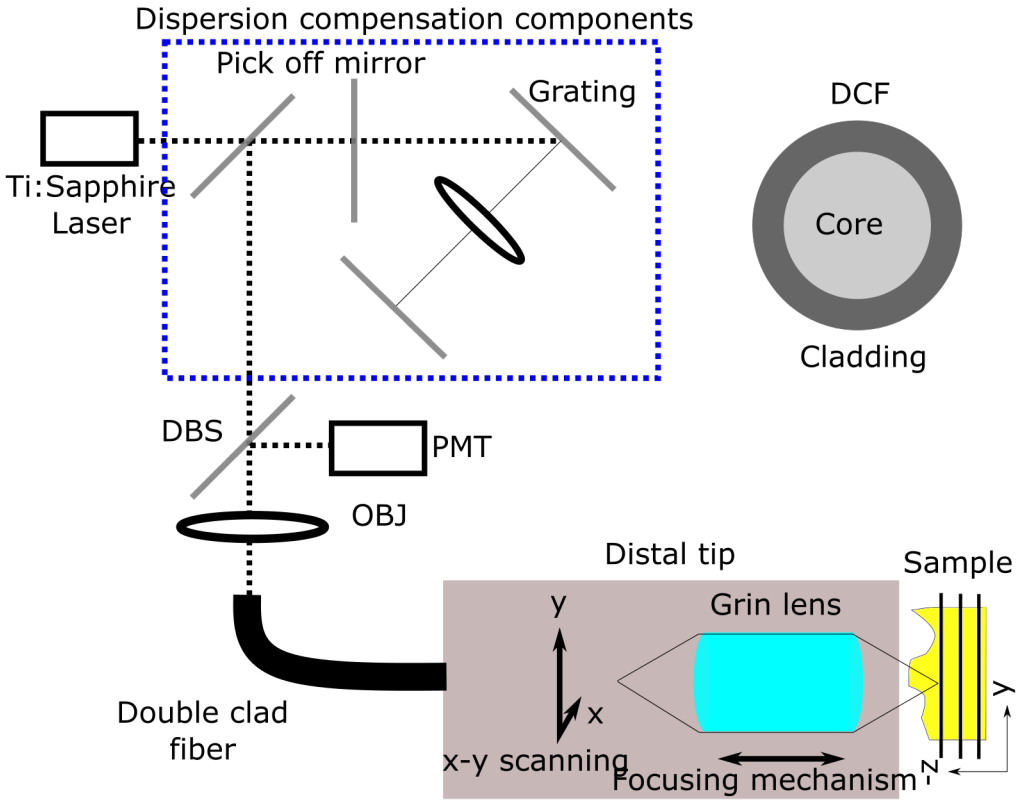


FIGURE 2.24: Two photon fluorescence endomicroscope.

Kim et al. demonstrated an endomicroscope compatible configuration based on a piezo element with a distal scanning fibre (2006)[106]. The distal tip has an outer diameter of 2.4 mm (Fig. 2.24). The core of the double-clad fibre is used to transmit emission light from the proximal to the distal end. The core and cladding are then both used for collection. Two-photon fluorescence has a crucial light collection advantage over single photon in that all light returning from the sample must have returned from focus and is collected. The photon

flux density requirements for two-photon fluorescence are only satisfied at the focal point, and further, confining the collection area does not provide any additional optical sectioning benefit [107]. In the figure, the region enclosed by the blue dashed line shows the dispersion compensation components, necessary to prevent pulse broadening in the fibre [105]. This configuration employed a spiral scanning pattern, while beyond the scope of this review, there is a fair amount of discussion about optical scan patterns in the two-photon literature [100, 105, 108] with a discussion about how to minimise: ringing artefacts, maximise sampling, and increase sampling uniformity, and do so in a way which is conducive to miniaturisation.

Other two photon endomicroscopes have been reported [109–111], one emerging example based on a micromirrors implementation. The micromirrors, which fall into a class of distally mounted microelectromechanical scanners (MEMS), are described by Piyawattanametha (2009) [104]. As of 2019, the smallest diameter that a DMD based endomicroscope can be miniaturised to is 5 mm. The rigid tip of that configuration makes it impractical for use in an endomicroscope.

2.7.7 Chip-on-tip

Chip-on-tip endomicroscopes feature the entire optical system including the camera at the distal end of the probe [112]. The system reported in 2017 by Matz, et al. featured an endoscope with the distal tip dimensions shown in Fig 2.25 [113]. The distal optics are shown in Fig 2.26. Incident light is guided from a single core multi-mode fibre into a small micro prism ($\sim$ 180 μm). That light reflects off a dichroic beam splitter and through an objective onto the fluorescent sample. The

fluorescence light then passes through the objective, transmits through the beam splitter and is imaged onto a miniaturised camera via a GRIN lens. Not shown in the diagram is a sapphire window just before the sample. Sapphire crystal is known for its resistance to scratches and to cracks [114]. The field of view is 160 μm. The system is capable of a lateral resolution of 0.74 μm, as well as a speed of 30 fps.

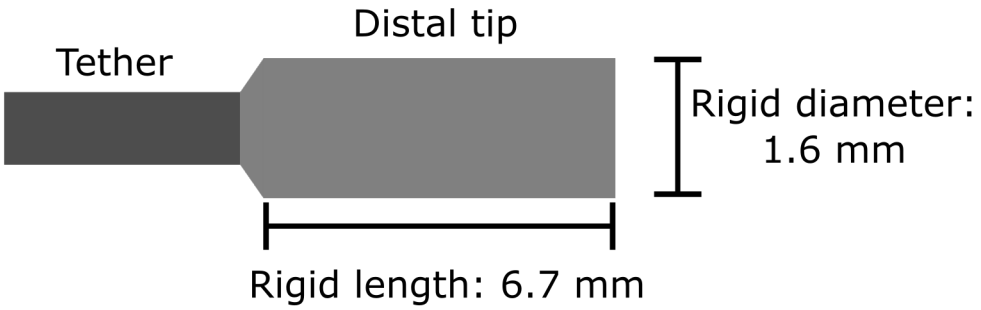


FIGURE 2.25: Rigid tip dimensions of a chip-on-tip probe.

There are always tradeoffs during miniaturisation of elements in the distal tips, and this emerging design appears to balance them well. Several disadvantages to this current approach are the lack of optical sectioning, and inability to transmit the light-field. Optical sectioning necessarily requires scanning or a structured illumination-based approach which implementing in this configuration would require miniaturised scanning elements, or some way to transmit patterns to the distal end. Additionally, fibre bundles do not just transmit intensity; they also transmit information about the phase of light, encoded in intensity. The spatial information can be descrambled to produce 3D information about the sample [90, 115]. This information is not preserved in this approach. One key advantage to chip-on-tip is the fact that it is not at all sensitive to bending when traversing an endoscope. It also is non-scanning, meaning there are no mechanical elements in the distal tip. Overall, this emerging approach appears to deliver

good resolution, and speed, all with very acceptable distal tip constraints. If axial sectioning was to be introduced, this appears to be an emerging rival candidate for fibre bundle-based techniques.

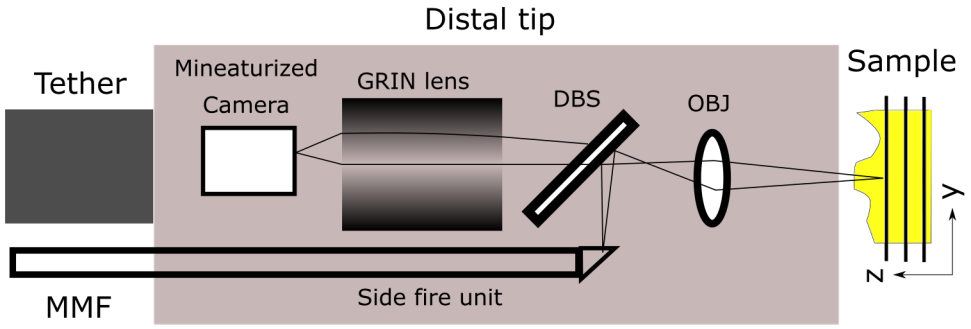


FIGURE 2.26: Chip on tip endomicroscope probe. Light is guided to the sample via a side-fire multimode fibre, fluorescence excitation is filtered and imaged onto a camera via a DBS and GRIN lens.

2.7.8 Phase contrast

Ford, T et al. reported a transmission contrast endomicroscope in 2012 [116]. The device, a phase contrast technique similar to differential illumination contrast reported in various forms [7, 61, 62], is based on the deviation of oblique rays through a sample. The endomicroscope proposed, as shown in fig 2.27, is unique in that the primary source of illumination is transillumination. However, the illumination source and detection optics are in reflection geometry, this has been previously demonstrated in bench-top systems [117]. Scattering allows the light emanating from the source fibres to act as an illumination source in the tissue. Two images are acquired, one with each illumination fibre active, the difference between the two images yeilds a phase contrast image. The probe is operated in contact mode which avoids collecting light which has reflected from the surface. The light reported from each distal fibre is 25 mW; in contrast, the Cellvizio

system has an approximate time-averaged power, during acquisition, of around 1-5 mW. The authors further report the use of a polarising beam splitter to reject the strong specular reflections. The approach that was used to attenuate the core pattern was identical to the iterative approach described in Sect. **??** and located in Refs. [86, 118, 119]. The images acquired when the left and right single fibre are illuminated have different intensities. A low pass filter with a large Gaussian kernel is used to normalise the images. Where $< I_i >$ refers to the image I_i after it has undergone a low pass filter the equation for the reconstructed image becomes:

$$I_{pc} = \frac{I_L}{< I_L >} - \frac{I_R}{< I_R >} \tag{2.9}$$

Ford et al. [116] ran a Monte Carlo simulation to illustrate that (1) light returning to the fibre optic probe does not return from a position close in the sample, and (2) the light which is generated by the fibre forms an effective transmission source position in the tissue. The stop in the distal end of the optics blocks light with a phase gradient. Phase contrast is a two-shot approach, and the authors used a double shutter camera for image acquisition.

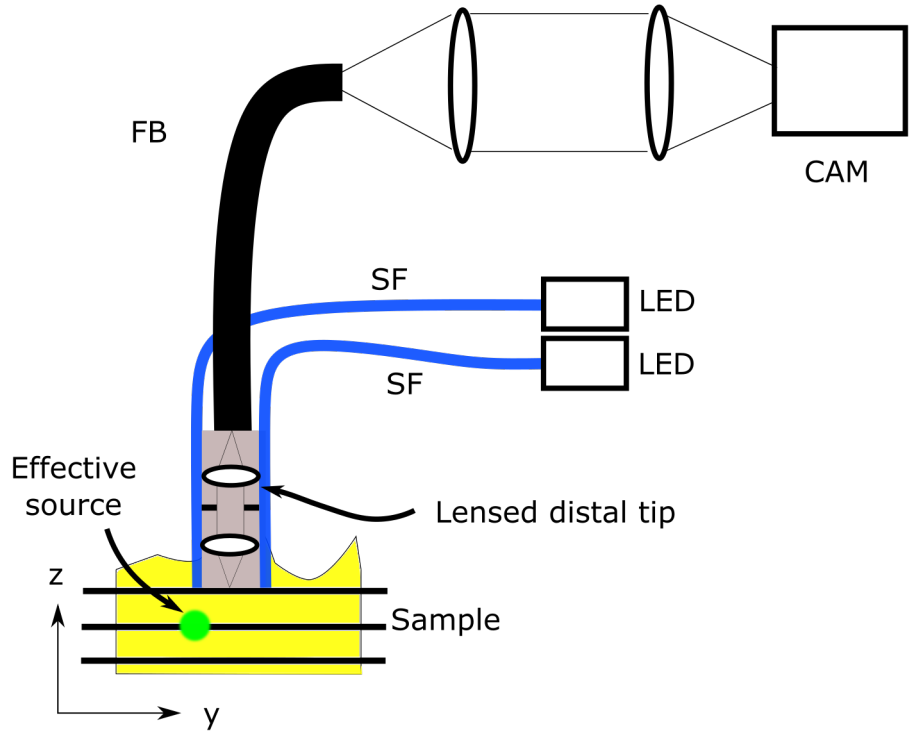


FIGURE 2.27: Phase contrast endomicroscope. Single fibres, outside the bundle, illuminate deep in tissue creating an effective source which is sensitive to phase contrast.

Chapter 3

Image processing methods

3.1 Introduction

Undesirable effects accompany imaging through a bundle. Core effects include the pseudo-honeycomb structure superimposed on the images and transmission properties that vary by core, such as the transmission efficiency and spatial distribution of light after being transmitted through a fibre. The latter can result from light coupling into fibre cores from different angles or effects due to the slightly different core geometries. Several approaches to cladding removal will be discussed: Gaussian filtering, median filtering, spatial frequency domain ring filtering, Delaunay triangulation, and local region thresholding. Background effects include ambient light degrading images, electronic noise in the camera, fluorescence generated in the fibre or reflections. Background removal is handled during processing; a robust procedure for background subtraction is developed in Sect. **??**, including removal of electronic noise, autofluorescence in the fibre, reflections, and room light, as well as compensation for transmission efficiency. Each of the following chapters uses components of the more comprehensive background subtraction presented in the Sect. 3.3.

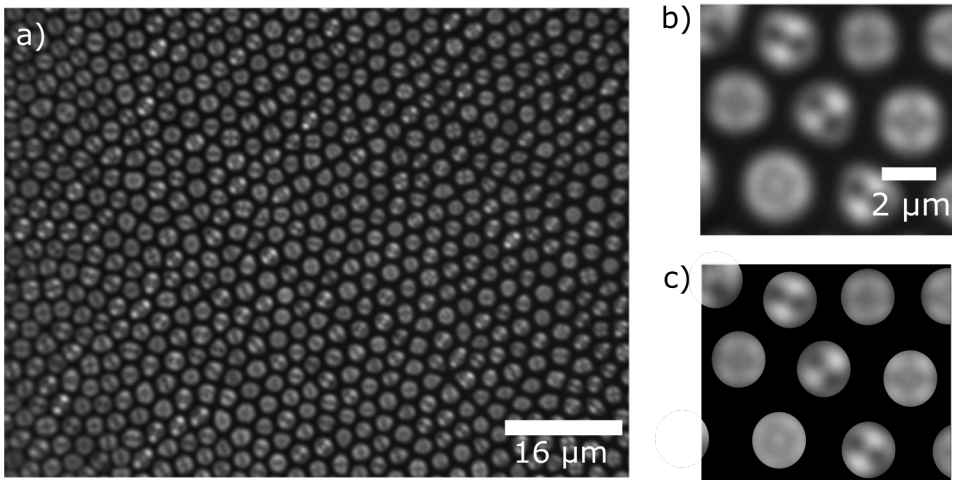


FIGURE 3.1: (a) Image of a distally illuminated 620 µm, 30,000 core fibre imaging bundle. (b) 2.5x digitally magnified image, (c) isolated cores of magnified image.

Fibre optic bundles are typically used as probes because they are small ($\sim$ 1 mm) allowing them to be used in the working channel of an endoscope or in a standalone mode. Typical histology sections can be around 3 mm x 3 mm. The area is around an order of magnitude greater than the fibre bundle. It can be useful to increase the clinically relevant information by translating the probe across tissue and stitching together subsequently acquired images in a process called mosaicking. Mosaicking schemes can vary in complexity from computationally expensive approaches to correct for scene deformation [16] and cumulative registration error [15, 120] to simple yet fast schemes that are suitable for real-time operation [16, 121]. When sampling a bundle the resolution is limited by the inter-core spacing of the bundle, by merging images with sub-core shifts, approaches to improve resolution through mosaicking have previously been reported [14, 88, 122]. Regardless of the scheme used, there are a few similarities, sequentially acquired images need to be registered, then fused into an existing mosaic. Scene deformation can be caused by several things, such as tissue deformation, motion

artefacts either in the sample, during camera acquisition, or as part of a multi-frame process. Following cladding artefact suppression, mosaicking is discussed (Sect. 3.4).

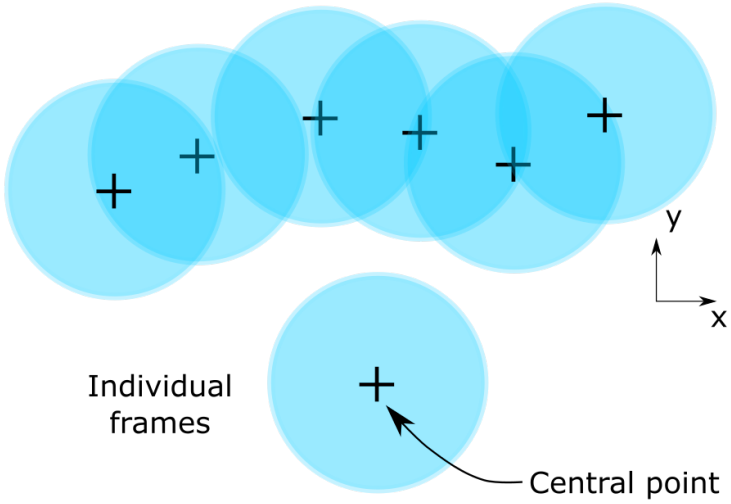


FIGURE 3.2: Principle of rigid translation. Subsequent frames are template matched, then fused into a mosaic.

The effectiveness of a cladding artefact suppression routine, to remove the structure from the images, can be benchmarked by how well five conditions are addressed. The five conditions are:

1. Remove cladding
2. Maintain resolution
3. Maintain contrast
4. Do not amplify out-of-focus light relative to in-focus
5. Speed must allow for real-time operation

The fabrication process of a fibre bundle involves stacking glass rods into a hexagonal pattern then drawn through a heating element until the fibre becomes

very thin. The drawing process causes deformation in the periodicity of the structure, this introduces variation in spacing between cores -affecting resolution, and the shape of each core - affecting transmission properties such as throughput and core-to-core coupling.

The Fourier space representation of a bundle sampled near the Nyquist limit is shown in Fig 3.3(a,b). The core structure appears as a ring. A radial profile is shown in Fig. 3.3(c), where the peak corresponding to the cladding is identified. Since the pattern forms a ring, the cores are not spaced uniformly enough for spatial filtering to work by cropping pinpoint locations, and the non-uniformity in core spacing introduces a broadening to the ring. In typical fibre bundle imaging, the proximal face of the bundle contains several pieces of information. It has (1) the intensity of each core, (2) a cladding structure, and encoded in each core is the (3) phase information of the light incident on the distal face. Both the cladding and spatial encoding of phase information must be discarded while maintaining resolution.

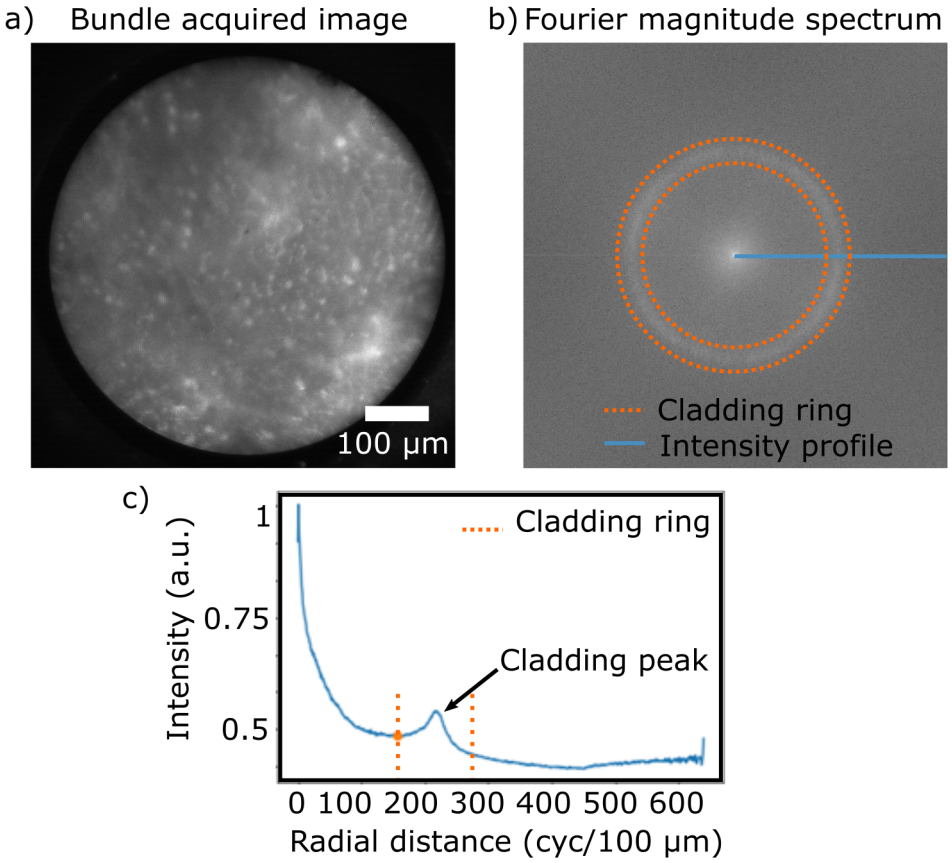


FIGURE 3.3: (a) Widefield image of bovine stomach acquired with a 620 µm probe. (b) Fourier domain representation. (c) Radial intensity profile identifying cladding peak.

Images acquired with a fibre bundle probe have a cladding structure superimposed. The cladding needs to be removed so that images represent the ground truth more faithfully. Furthermore, when images are used for mosaicking, the shift is often registered in the spatial frequency domain. It is problematic to do this with raw bundle images since the magnitude spectrum of the image containing a pseudohexagonal pattern has substantial energy concentrated in the cladding ring.

3.1.1 Gaussian filter

A simple way to accomplish cladding artefact suppression is through the convolution of the image with the discrete approximation of a Gaussian using a 2D Gaussian kernel. The size of the kernel needs to be selected to balance a few factors including computation speed, blurring of neighbouring pixels, and complete cladding removal [123]. An early source was Aslan et al. (1999) [124] who accomplished this with a modification to the optical system which performs the same operations as a Gaussian smoothing filter. Aslan et al. used commercially available "fibrescope filters" which expand the image of each fibre. The paper surveyed clinicians and reported positive feedback when core smoothing was applied. High-speed Fourier domain operations have made post-processing operations, which do not require modifications to the optical system, more common [125].

The effect of a Gaussian filter as a function of its width is now shown. The kernel width is reported as the ratio of the Gaussian width to the inter-core spacing. To determine the magnification from the bundle plane to the camera Fig. 3.4 was used. The plot was obtained by drawing a line across several cores and plotting an intensity profile. In pixel space, the image shown above has an inter-core spacing of 45.6 pixels which is equal to 3.2 μm. The non-integer pixel value is due to averaging over five cores. Each pixel is equal to 70.2 nm at the object plane. With a pixel spacing of 2.5 μm, this implies a magnification of 35x. The intensity of modal information can be seen near the peaks; as a result, local minima were used to avoid ambiguity.

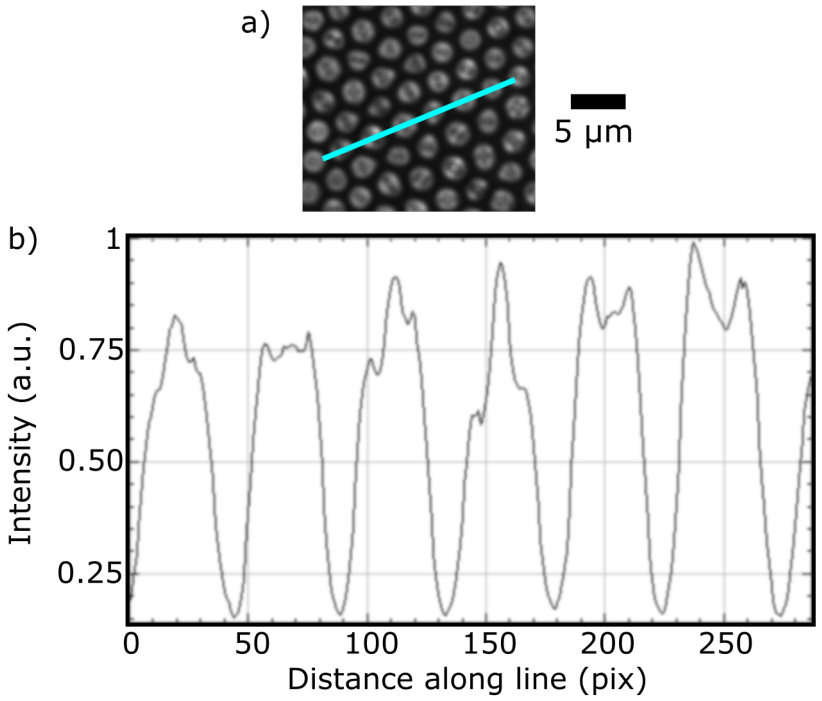


FIGURE 3.4: (a) Image of proximal face of a 620 µm fibre bundle, illuminated distally by a cell phone LED, (b) intensity profile across 6 cores.

The image to be smoothed was acquired using a fibre bundle, which was illuminated with a mobile phone LED. Applying a Gaussian filter with different kernels yields the cladding removed images shown in Fig 3.5. Each discrete Gaussian kernel is area normalised, leading to an equal intensity in the processed and pre-processed images. At the maximum Gaussian kernel width tested, the intensity has increased negligibly by 0.38%. The figure seems to imply reasonably good cladding removal with a kernel width of at least 1.5 $\times$ inter-core spacing (ICS). It might be tempting to argue that the non-uniform intensity in these images suggests a higher width is needed, but no correction has been made for the coupling efficiencies of the cores.

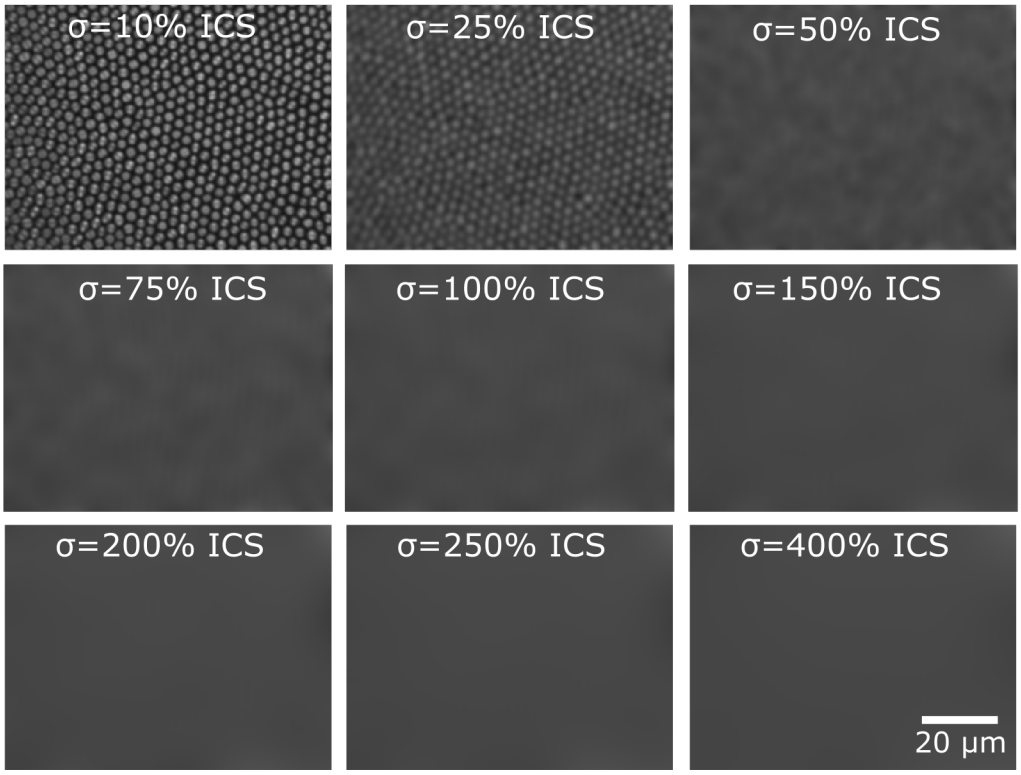


FIGURE 3.5: Gaussian filtering of reference image (Fig. 3.1(a). Filtering kernel displayed relative to mean intercore spacing (ICS = 3.2 μm). σis the width of the Gaussian kernel.

The Gaussian kernel is a discrete 2D approximation of a continuous Gaussian function corresponding to a certain width. Its half-width half max is represented by σ. As part of a well-known image processing step, a blurred image is given by convolving the discrete kernel with a raw image. Fig. 3.5 shows the convolution of different discrete Gaussian filters and provides qualitative information about minimum kernel widths needed to remove artefacts from an image. Although 400% appears to completely remove the cladding, this does not make any claims about how the resolution is maintained, and a lower percentage is often sufficient. Under the Gaussian convolution, each core is blurred over an area not just to cover the cladding but to cover regions sampled by other cores as well. The blurring operation means that contrast is not strictly maintained with this

approach; for this reason, Gaussian filtering is not considered an overly quantitative approach. An advantage to Gaussian convolution is that processing can be done in Fourier space, significantly reducing computation time (2 fast-Fourier transforms (FFT), 1 pointwise multiplication).

As part of a Gaussian blur operation it is common to use a normalised Gaussian kernel to preserve intensity under the blurring operation. In conventional fluorescence imaging this would maintain quantitative fluorescence information. That however, is not true in bundle imaging. Light returning from the sample to the bundle which is incident on the distal cladding is lost by the optical system, not guided into the cores. Consequently, only core transmission reliably corresponds to sample intensity. A more accurate approach is, during processing, to increase the intensity of the cladding pixels rises to the level of nearby cores, a computationally efficient version of this approach is later presented in Sect. 3.2 A simple geometric correction for fill factor, the ratio of core area to total bundle area on an image, would also account for the intensity drop. Under the Gaussian operation, the spatial distribution of intensity within the cores is also lost, although this is often desirable since the core distribution itself is not the same as the features on the distal end; this requires complex processing to descramble.

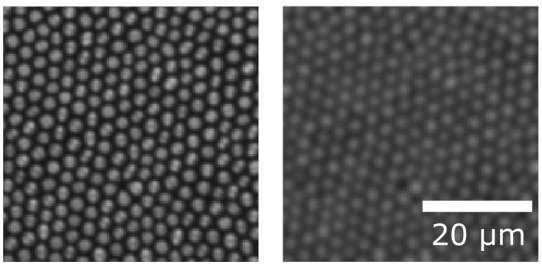


FIGURE 3.6: Blurred image (Fig. 3.1(a) showing suppression of modal information

How well resolution is preserved was studied by processing a resolution target (Fig 3.7). The Gaussian kernel width was chosen in Fourier space to be 1/2 the spatial frequency of the cladding. A radial intensity line profile of the resulting magnitude spectrum of the processed image shows the virtual elimination of the cladding peak. The minimum resolvable feature similarity in the two images shows that resolution is maintained under Gaussian processing.

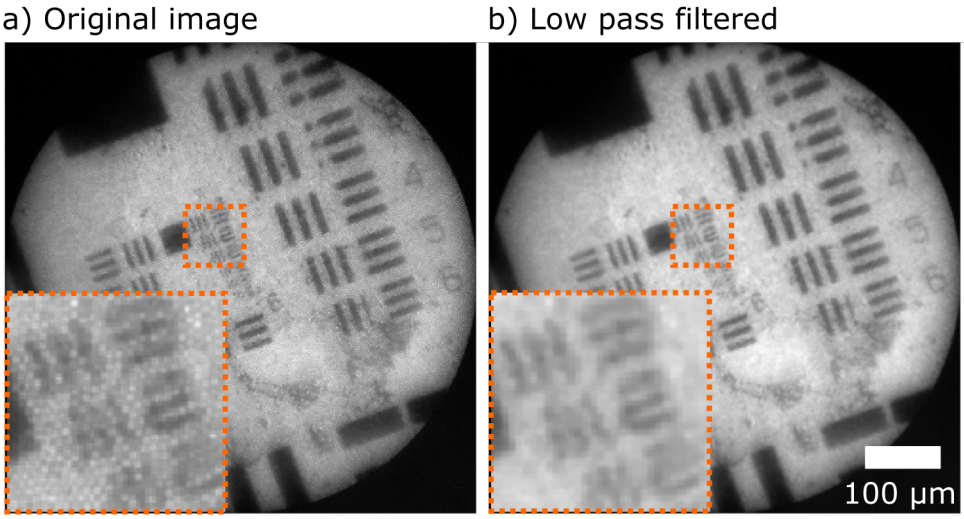


FIGURE 3.7: USAF resolution target before and after processing with a Gaussian width 50% intercore spacing.

In summary, a Gaussian filter presents a computationally inexpensive way to remove the cladding structure, which mostly addresses the five conditions. The cladding is removed with the minimum Gaussian width on the order of 1.5 $\times$ ICS. Resolution is mostly maintained. Contrast is mostly maintained, but not entirely, the nature of the Gaussian is to blur the image beyond the cladding boundaries. Low contrast applications that require sampling near the resolution of the bundle may need a more robust method. Out-of-focus, light is not amplified by this method. Finally, speed is the most significant advantage, as the Gaussian convolution operations can be performed in the Fourier domain.

3.1.2 Median filter

The median filter involves a square window which "scans" across an image and arranges the pixels in that window by intensity from increasing to decreasing. The median filter consists of the selection of the median value, in the case of Fig 3.8, that value is 7760, also the central value in the 5x5 window. The central intensity value is then inserted as the pixel value in the processed image. The filter is an ordering and selection filter (OSFILT). In the programming suite, LabVIEW, an OSFILT is referred to as an "nth-order filter" since it orders the pixels and selects the nth pixel. This filter is robust to outliers which will fall at the extrema of intensities. This sweeping operation is computationally expensive, and it cannot be performed in the Fourier domain.

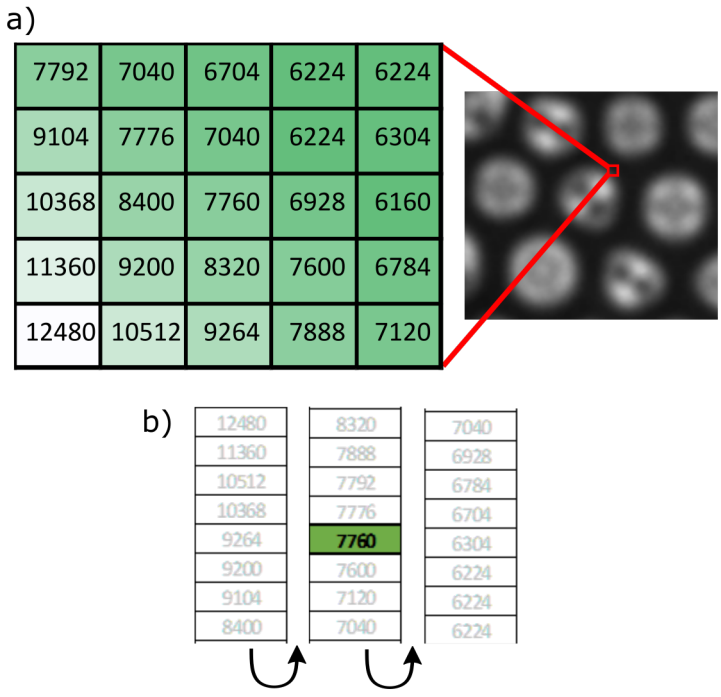


FIGURE 3.8: Process of median filtering an image to remove cladding artefacts. (a) Filter window scanning across image of core structure, (b) values arranged by intensity and central value selected.

An image with an inter-core spacing, in camera pixel space, of ~ 3 pixels was

processed with a median filter of 3x3 pixels. The maximum resolvable USAF element was group 7 element 2 in both images. With the exception of clear blurring introduced resolution (condition number 2) appears mostly maintained (Fig. 3.9).

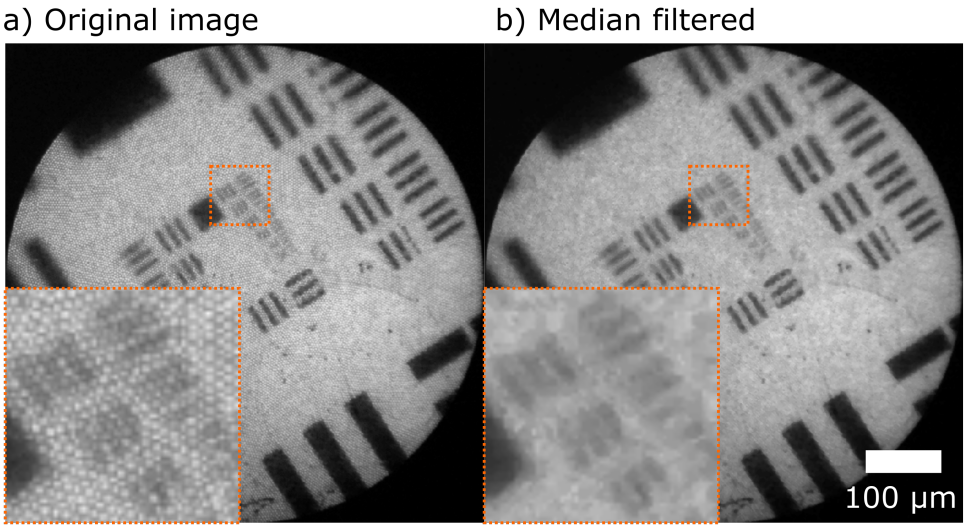


FIGURE 3.9: Raw image before and after processing with a 3x3 µm median filter. Median filter kernel width approximately equal to inter-core spacing.

The median pixel is not always the optimal selection. To use the median value, the window must be large enough such that when the window is sliding across a dense cladding region, the median pixel must always be a core pixel. Since the cladding thickness is less than the core size, the size criterion means intensity values from neighbouring cores are used to determine pixel values of a single core, or intensity is bleeding into neighbouring cores, diminishing contrast (Condition number 3). There is a case to select the highest intensity pixel. The reason for this is that it is not possible to have a smaller widow which only needs to cover the cladding and a single-core pixel. The small window case might be desired to limit costs in terms of low contrast. Whether or not this is suitable depends on the sampling of the bundle. A bundle sampled with few CCD pixels per core

area means the spatial intensity distribution in each core has mostly been averaged over, and intensity is approximately equal at the edges and centre, in which case this will work well.

The median filter is not optimised when camera pixels resolve finer detail in each core as opposed to simply averaging the intensity. The critical result of work presented by Orth, A et al. (2017) [89] was that the central part of a core encodes light paraxially incident, and light incident from an angle was "pushed" to the edges. From a practical perspective, all intensity at the edges ought to be removed by a filter, attenuated, or at least not amplified relative to light from the core's centres. Choosing a window that only covers the cladding and only slightly overlaps a single-cores edge means that the core pixel chosen will be from out-of-focus light. While a smaller window can lead to maintaining resolution, cladding regions will be replaced by intensity at the edges, failing condition 4.

There are a few considerations when selecting OSFILT parameters which all directly correspond to the five conditions. First, to remove the cladding, the median filter window must take up a large enough area so that the median pixel value is always in a fibre core. Second, to maintain resolution, the OSFILT window should be as small as practically satisfies full cladding removal. Third, the OSFILT is a non-linear process, and intensity is not strictly maintained; however, it is bounded by the values in the pixel window. Fourth, the OSFILT must not amplify out-of-focus light. The median filter does not, however choosing the brightest pixel can lead to the amplification of light at core edges. Fifth, to limit speed, the window chosen must be small enough to minimise the number of operations.

3.1.3 Spatial frequency domain filtering

Spatial frequency domain filtering or disk filtering was first reported for a fibrescope by Dickens, M in 1998 [126]. It is based on the principle that the cores form a clear band in the Fourier domain, and without severely degrading image resolution, this band can be rejected, and the image reconstructed. A disk, corresponding to the thickness of the radial cladding peaks in the magnitude spectrum is removed from the Fourier transform, and then back-transformed (Fig. 3.10(a)). The filter appears to have filtered very effectively and as a Fourier domain operation, efficiently. The computation cost seems to be similar to Gaussian filtering (2 x FFT's, 1 pointwise multiplication). The location of the band in Fourier space only needs to be determined once and saved to memory. When sampling near the Nyquist limit, this approach appears to be an excellent high-speed candidate to remove the core and cladding patterns.

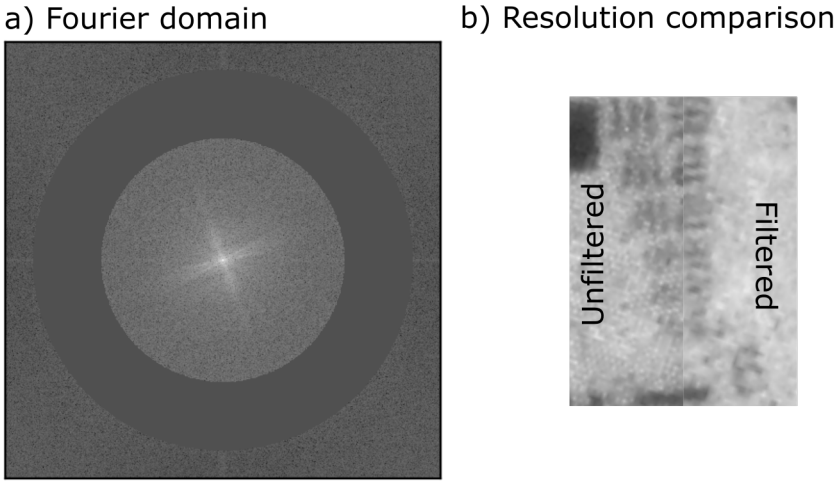


FIGURE 3.10: Cladding rejection via band reject in the spatial frequency domain. (a) cropped Fourier representation, (b) processed image comparison.

The high sampling case, when modal information can be seen in the cores, disk filtering requires an additional step. Consider an illuminated fibre in Fig 3.3(a). The Fourier spectrum that corresponds to this has several distinct bands.

These bands correspond to the modes in the fibres, as well as structural features such as sharp edges. The inner band corresponds to the fibre core structure.

Removal of only the inner band produces the image shown in Fig 3.11(c). It is clear from this that while the core structure is removed the modal information in each fibre that needs to be averaged over is not. The fact that the modal information in each core remains localised also hints that by disk filtering generally resolution is not impacted. This modal information, in most cases, needs to be discarded. By expanding the diameter of the annulus to include the higher spatial frequencies that correspond to the modal information, the image in Fig 3.11 is generated. It is clear from this image that the core, as well as the higher frequency modal information, is suppressed. It follows that if the resolution is preserved in the low sampling case, it must be maintained in the high sampling case, since all frequencies higher than the core spacing, are below the sampling frequency of the bundle.

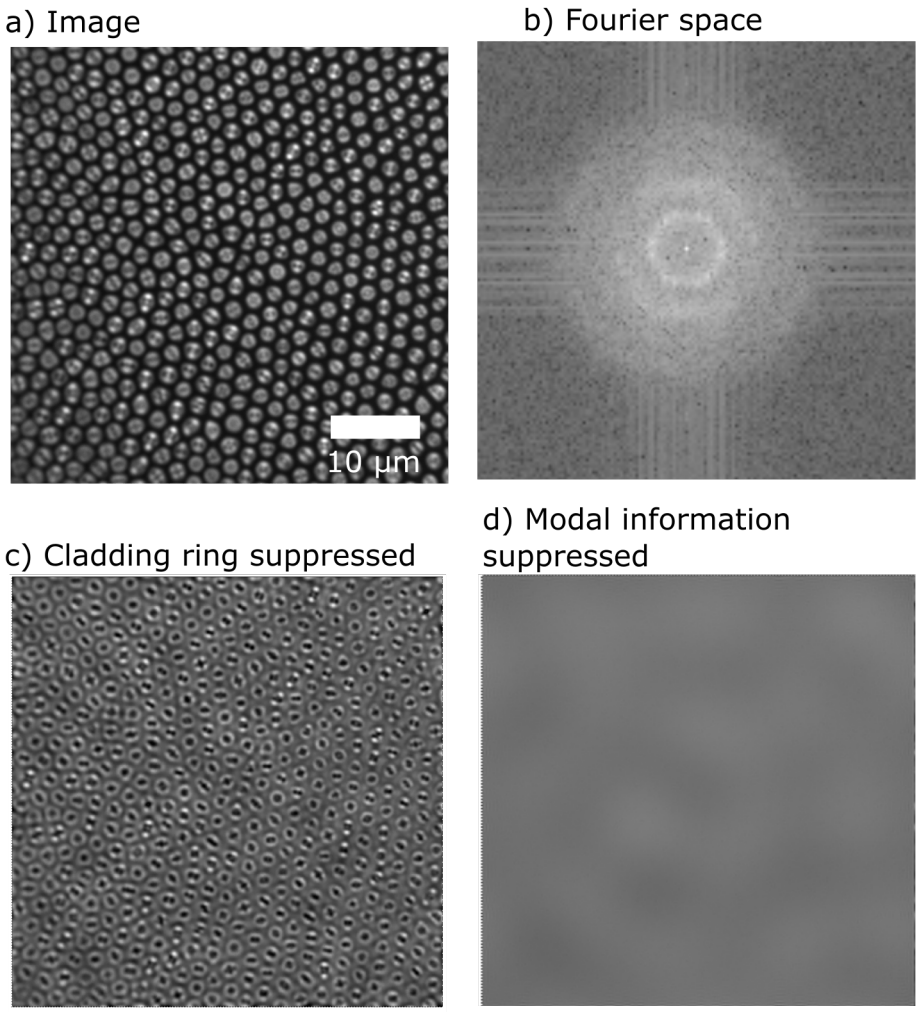


FIGURE 3.11: Spatial frequency domain filtering when bundle sampling includes modal information. (a) raw image, (b) Fourier space representation, image reconstructions with (c) only cladding ring suppressed, (d) suppression of cladding and higher frequency rings.

In summary, spatial frequency domain filtering removes the cladding, maintains resolution, and contrast, does not amplify out-of-focus light relative to in-focus. The annulus must include modal information embedded in each core. The speeds possible are more than sufficient for real-time operation. The only drawback is the calibration step where the location of the radial ring needs to be determined.

3.1.4 Delaunay Triangulation and interpolation

The approaches above were all relatively easy to implement and computationally efficient. There were drawbacks mainly in the loss of resolution and contrast. A more sophisticated class of cladding removing approaches involves core localisation and interpolation. The first such method of this class will be the Delaunay Triangulation. The triangulation method was reported in a conference proceeding by Elter et al. in 2006 [127]. Core localisation is first discussed and then triangulation.

The core localisation step involves systematically finding the central points that correspond to the $\sim 30,000$ cores in a multicore fibre optic bundle. First, a calibration image is acquired with a uniformly illuminated bundle. A 3D wireframe plot shows the intensity in an arbitrarily chosen local area (Fig. 3.12). A naive way to approach this may be to apply a simple threshold mask. The problem with a global threshold is the illumination profile across a bundle often varies by more than the peaks in a region. The results of a threshold mask are shown in Fig 3.13.

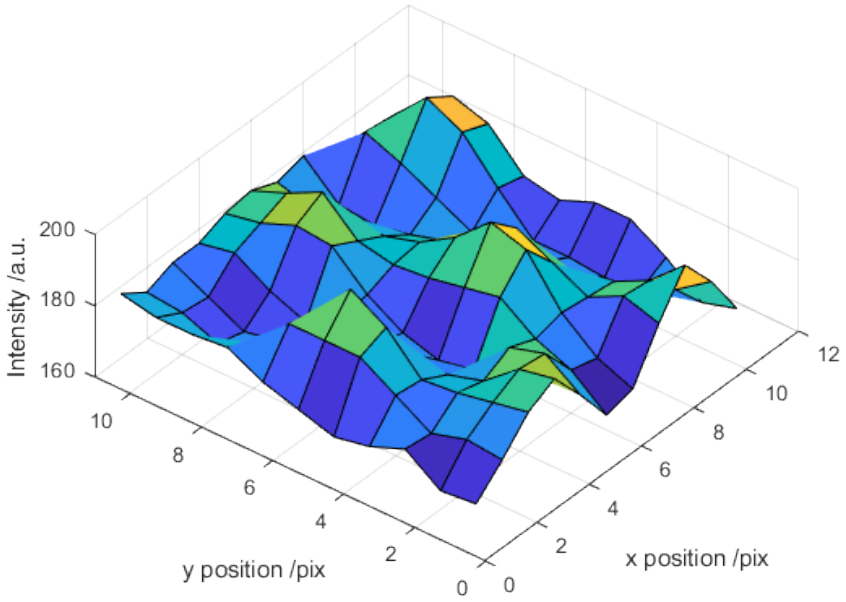


FIGURE 3.12: Wire plot of the intensity levels of a uniformly illuminated bundle.

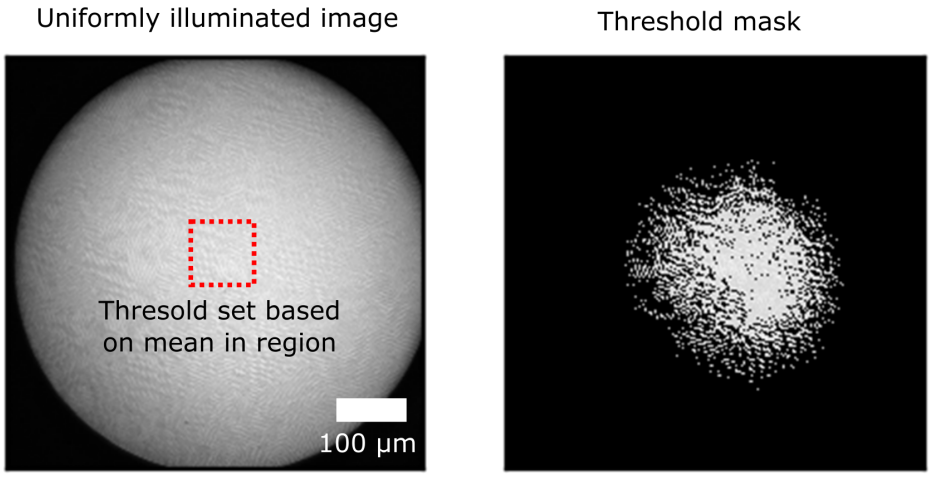


FIGURE 3.13: Threshold mask taken without correcting for bundle intensity profile. Threshold value set at mean value, taken over central 100x100 pixel square.

Elter, M. et al. proposed handling core finding in a local region as a way to account for global variation [127]. Ford, T. et al. did also show in 2012 [86] it is possible to divide by a low pass filter version of the original image and apply a global intensity mask. The steps presented by Ford have been reproduced (Fig

3.14) with the same threshold criteria as above. A local region is shown in the inset of Fig 3.13. It appears globally to correct for significant intensity differences, and locally seems to produce a relatively good first approximation of core locations. It does not entirely isolate cores to a single-pixel however. The criteria were slightly raised to determine if a more stringent threshold would accomplish core isolation; this was done by including only pixels with intensity values 5% above mean intensity. The stricter criteria led to a breakdown in the global profile with some very well resolved regions, and some which had vanished entirely.

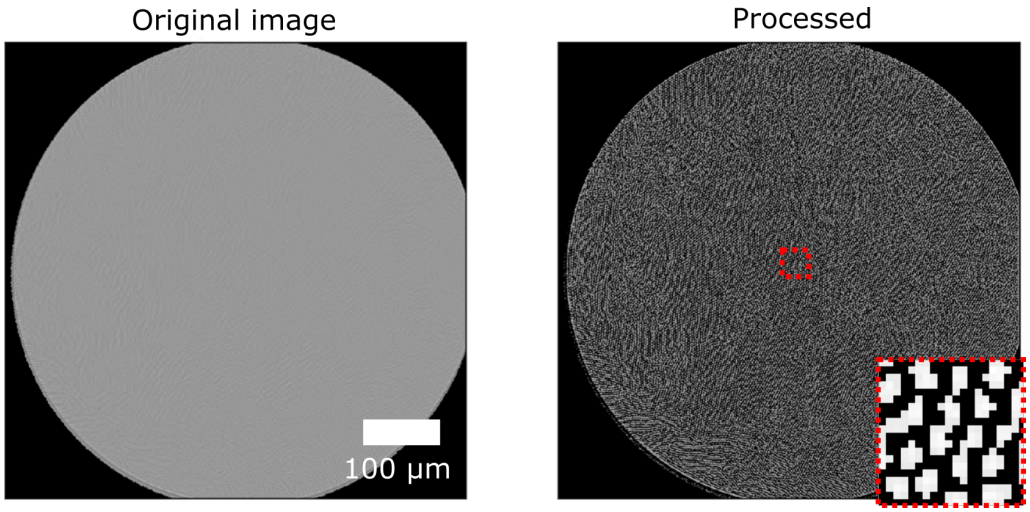


FIGURE 3.14: Improvement in pixel contrast after using local areas as threshold reference.

The threshold mask performance was compared to that of the iterative local approach Ford suggested. The first process in the iterative procedure was to identify candidate regions where the intensity difference exceeded a threshold. Then calculating the residuals squared of a Gaussian with constant kernel width set to the average core diameter. Then the residuals-squared was determined, and candidate points that exceeded a threshold were discarded. A local region of the original image is shown in an inset in Fig **??**(a), a window threshold mask is applied to isolate the clearest cores. The window threshold mask is only used

to evaluate the false-negative rate of the iterative approach, not as part of the localisation process. A parameter optimisation varying Gaussian kernel width and score threshold were attempted. The parameter optimisation was a mostly arbitrary process, and performance, which can be seen in the iterative approach results with clear missed cores, on the unnormalised image was poor.

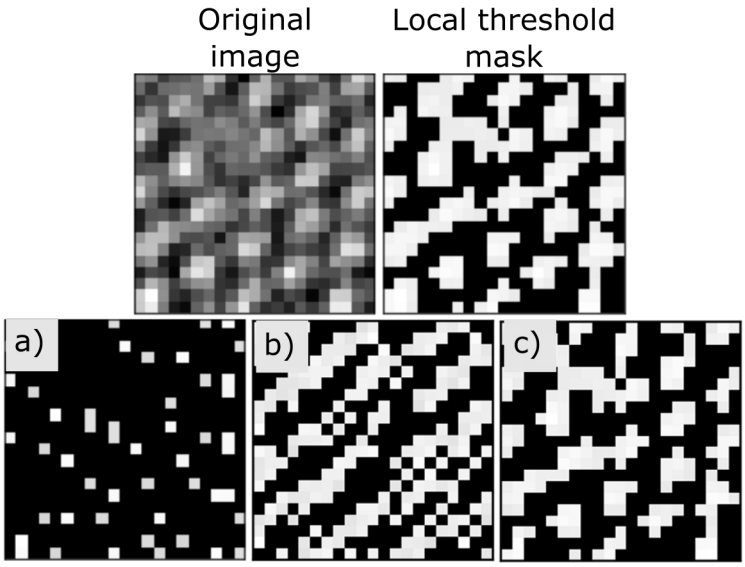


FIGURE 3.15: Cladding artefact supression via (a) local area comparison, (b) local area comparison of normalised image, (c) local area comparison with sub-pixel Gaussian fitting.

The localisation was repeated for the normalised image (Fig 3.15(b)), which delivered superior results in terms of eliminating false negatives. There does not appear to be a case to use the iterative approach to approximately localise the cores into clumps, without pre-processing. Taking a threshold mask of a normalised image considers global intensity variation and delivers subjectively better first approximation performance. Additional improvement can be seen when the threshold masked image is iteratively cycled with the Gaussian fitting (Fig 3.15(c)).

It is clear from Fig 3.15(b) that the cores form clumps, individual pixels must be isolated since the process of triangulation requires a point map. The authors continue the process of fine selection by iterating through all non-zero pixels and fit a symmetric Gaussian centred on that pixel to the pixel creating another score array. The standard formula for a 2D non-symmetric Gaussian was used with free parameters corresponding to amplitude, x-width, y width, and mean. The 2D fitting leads to subpixel core positioning. On a refined mesh, this led to a revised match score for each pixel candidate. The curve fitting method used was Levenberg-Marquardt which was effectively a least-squares method that is slower than algorithms like Gauss-Newton, or gradient descent, but needs a less precise initial guess.

Finally, candidate points are swept in the order of highest score to lowest based on score only. The highest scoring pixel is assigned as a core centre, and a distance constraint is applied, and only pixels outside the radius of a core centre can be assigned themselves as core centres. Fitting is done in high precision space, meaning it is unlikely that pixel limits of the image will adversely contribute to this constraint. The ultimate result is a point map with the central pixel in each core selected.

The process of core localisation is separated into two distinct parts: separate core regions into clumps, isolate centre through the non-linear fitting of a Gaussian. The fitting leads to a point map with sub-pixel resolution and prepared for the step of interpolation.

Triangulation is a process of interpolation connecting points by triangles. It is derived from computational geometry; this is not a comprehensive review but

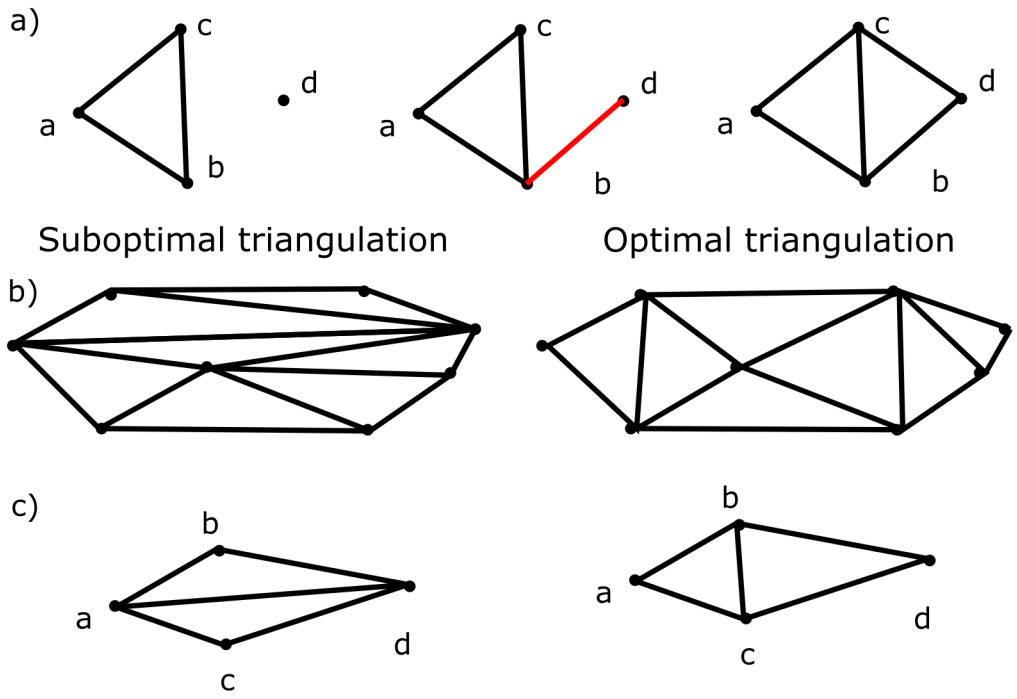


FIGURE 3.16: (a) Process of connecting points via triangulation, (b) qualitative comparison of optimal and suboptimal triangulation, (c) edge flipping to optimise triangulation.

does introduce a few principles. The basics revolve around how to connect points in a plane. The idea is any line segment drawn from one point to another must complete a triangle; it also cannot intersect another line segment. Consider Fig 3.16(a). If ABC is connected, and point d is inserted into the point-map, connecting only one line segment means the plane CBD remains open, this is not an example of triangulation. Connecting CDB however, leads to a closed shape which is a triangulation.

The goal is to have triangles in as local of a region as possible. For instance, it would not make sense for a triangle to cover the length of a bundle. The same point map is shown on the left and right-hand sides of Fig. **??**(b) with two different triangulations, one suboptimal and the other optimal. Visually it is clear that the triangles are more localised in the second example, and less so in the first.

While beyond the scope of this, there is a crucial advantage in processing power with more localised triangles [128]. One way to say the triangulation is optimal precisely is the smallest angle in each triangle must be as small as possible. With a set of sparsely distributed points and two different triangulations, quality can be determined by comparing the smallest angles. If the sum of the smallest angles in one set of triangulations is larger than another than that triangulation is more optimal. The simplest way to optimise this to carry out any triangulation scheme, then follow easy rules algorithmically that can optimise a triangulation. One such example of optimisation is edge flipping, shown in Fig. **??**(c).

Delaunay triangulation is an optimised solution of a triangulation. Voronoi Tessellations are the dual graph to the Delaunay triangulation. A Delaunay triangulation involves each point in a point map on a vertex of a triangle. In a Voronoi diagram, each point sits at the centre of a plane. The plane is then designed, so all locations in the plane have a Euclidean distance closer to that point than any other. The interpolation is created using a natural neighbours scheme. Once a Voronoi Tessellation has been completed the pixels of the camera are mapped to the Veronoi tessellation, and natural neighbours is a way of interpolating intermediate points.

The location of the cores with high precision is then determined, and so is the interpolation scheme. The image reconstruction step consists of fitting the predetermined non-linear Gaussian in two dimensions. Except, in this case, the values of x-mean, y-mean, x-width, y-width all remain constant. The only variable to optimise is a-amplitude. Once this is done, the point map is then interpolated using the predetermined scheme in real-time.

It is now possible to evaluate the approach in terms of the five conditions. The cladding is removed effectively since only the points corresponding to cores are interpolated. Resolution is also maintained. By the very nature of the Voronoi tessellation, contrast is maintained as well. At least with sampling near Nyquist out-of-focus noise is not amplified. The approach described here does not examine fitting when modal information appears in the cores. The majority of the computationally expensive steps occur as a calibration step leading to real-time processing. However, it still involves iteratively fitting Gaussians in a for loop sequence which is much slower than any of the other methods discussed.

3.2 Local area comparison

In 2012 Ford et al. developed a cladding removal algorithm which must be applied to each acquired image before HiLo processing [86]. Ford acknowledges that the localisation methods and interpolation using the Voronoi diagram are the most comprehensive; however, they are also the most time consuming; the following is proposed as an alternative.

The algorithm developed is:

1. Acquire a raw image of a uniform fluorescent surface, I_{raw}

2. Apply a Gaussian LPF to the raw image to generate a cladding artefact suppressed image I_{cs}, then take an additional low pass filter to generate LP(I_{cs})

3. Iterative loop: Sweep window across the image and compare pixel values in I_{raw} to the LP(I_{cs}) in the same region. Set a mask value to 1 if the intensity is greater or 0 if it is lower.

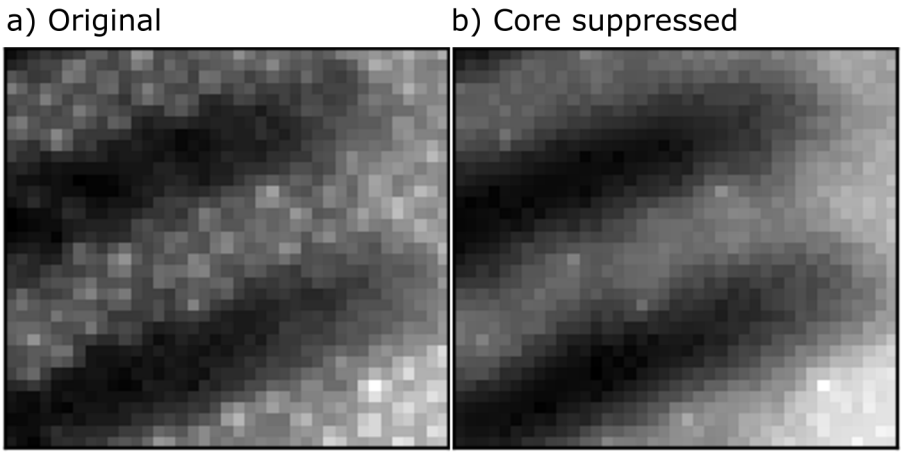


FIGURE 3.17: (a) Original USAF target image with superimposed cladding structure, (b) 2 iterations of local area comparison filtering.

After the loop is complete, generate a processed image with:

$$I_{csprocessed} = I_{cs} \cdot I_{raw} + (1 \breve{} mask) * \mathrm{LP}(I_{cs}) \tag{3.1}$$

The algorithm constantly maintains the intensity values of the centres of the cores and forcing the neighbouring cladding pixels to rise to those values. The results after 2 iteration is shown in Fig 3.17. Using the difference in the old and new cladding artefact suppressed images, the image reached a minimum difference at about 2 iterations. Resolution was also maintained since core pixel values are greater than the cladding pixel values.

Local area comparison filtering requires 1 array-wise comparison, 2 array multiplications, and one array subtraction. There were some major computational benefits to this approach; namely, there was no need for pointwise operations. It was also very simple to code and execute. A comparison of resolution appears in Fig 3.18.

Regarding the five conditions, the cladding is clearly removed, the resolution

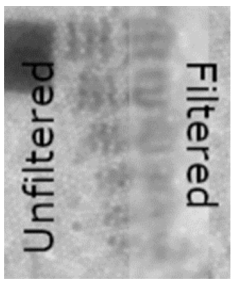


FIGURE 3.18: Local area comparison filtering resolution comparison.

appears mostly maintained, this is a natural property of the cores always been brighter than the local area, the downside is it also necessarily requires this. One could imagine a dark bead taking up only a single core on a bright background; if the cladding pixels were brighter the process would incorrectly remove this as cladding. Contrast is also mostly maintained; this needs to be done during the masking process. The process consists of keeping identified core pixels and adding in the LPF for those not identified as cores. The low pass filter needs to be weighted with a fill-factor correction. As previously discussed, the low pass filter blurs bright cores across an area, and intensity drops; the proportion this drops is the fill factor, i.e. the core area relative to the cladding. This approach, similar to other approaches, does not amplify out-of-focus noise. It also does not suppress it. The algorithms can be executed faster than the localisation and interpolation approach; however, it is much slower than the Gaussian alone, or median filter.

3.3 Background removal

A systematic way to generate background-subtracted images was developed by Ford et al. (2012) [86]. The authors report a way to correct for: the ambient background, the autofluorescence of the fibre, and the core efficiency. The background

subtraction process, described in an operational sense is to process the raw images using the local area comparison approach, this fills in the cladding pixels around the cores. Next, acquire three images: I_{bg}, I_{af}, I_{hf}. I_{bg} is an image acquired with the laser off, and is meant to account for ambient light and camera shot noise. I_{af} is an image acquired with the laser on and the probe in a dark tube filled with some non-fluorescent material, it is an image containing only the autofluorescence generated in the fibre. Finally, I_{hf} is an image acquired with the probe dipped in a homogenously labelled fluorescent fluid. I_{hf}, assuming no autofluorescence, background, and ideal core efficiency would be an array of pixels of identical values. To correct for the different pixel values, each raw image needs to be scaled by the inverse of a weighting factor W, where

$$W^{-1} = (I_{hf} - I_{bg}) - (I_{af} - I_{bg})\frac{P_{hf}}{P_{af}} \tag{3.2}$$

The first term is the uniform fluorescence with the background-subtracted. The second term is the background removed from the autofluorescence image. The second term is scaled by the relative power of the fluorescent background, to the autofluorescence generated in the fibre (P_{hf}/P_{af}). Once collected this weighting factor, corrects for core transmission coefficients, and is maintained during the whole image acquisition process. The processed image, corrected for transmission efficiency is now given by:

$$I_{processed} = W(I_{cs} - (I_{af} - I_{bg})\frac{P}{P_{af}} - I_{bg}) \tag{3.3}$$

where I_{cs} is the raw image after it has undergone the iterative process to suppress

the cladding. The total number of operations is 2 pointwise multiplications, and 4 point wise additions, this is of course in addition to the point wise operations required for cladding artefact suppression. As a matter of practicality, to save computation time, it appears useful to store:

$$I_{cal} = (I_{af} - I_{bg})\frac{P}{P_{af}} - I_{bg} \tag{3.4}$$

The calibration image can now be stored for use during the entire acquisition, as long as the probe doesn't shift, and laser power remains constant. Probe shifts lead to misregistration, and laser power leads to a change in autofluorescence. The reduced expression for the real time background subtraction is:

$$I_{processed} = W(I_{cs} - I_{cal}) \tag{3.5}$$

The number of real-time operations becomes 1 pointwise addition and 1 pointwise multiplication.

3.4 Real time mosaicking

The methods used in Chaps. 4, 7, 5 all involve real-time mosaicking with rigid registration and images which are then placed "dead leaf" into mosaics. The dead leaf approach is one where pixel values of the current frame overwrite values of the mosaic without any blending. These methods are described below.

3.4.1 Rigid registration

Real-time mosaicking often involves registering images using rigid-translation. Rigid translation can be determined using fast normalised cross-correlation (NCC). Additionally, correlation and matching are operations that can be used as a measure of similarity between two images. These operations are commonly used in mosaicking, and image processing generally. This subsection will introduce what NCC is and what its benefits are compared to correlation.

Correlation is used to extract the similarity between a reference image and a comparison image. Matching is performed by taking an inner product of two images. When comparing two different images to a reference, the pair with the highest inner product is the most similar. Correlation can be done between two full images, or it can be done as a sliding window operation. In a sliding window operation, a square is extracted from a reference image, a sliding window of equal dimensions then translates across the comparison image, at each location of the reference square a correlation matrix can be built up which contains the inner product of the reference square and the sliding window at each location of the sliding window. The formula for correlation is given in Eqn. 3.6

$$C_M = \sum_{i=-N}^{N} \sum_{j=-M}^{M} R_{sq}(j,i) I(x+j, y+i) \tag{3.6}$$

where R_{sq} is the square extracted from a reference image, I is the comparison image, and C_M is the correlation matrix which is computed at each position of the sliding window. This formula is useful for matching, however, Eqn. 3.7 below,

may be preferred.

$$\text{LSM} = \sum_{i=-N}^{N} \sum_{j=-M}^{M} (R_{sq}(j,i) - I(x+j, y+i))^2 \tag{3.7}$$

where LSM is the least-squares matrix. To perform a normalised cross-correlation, beginning with Eqn. 3.6, subtract the mean from each component, $\overline{R} = R - R_{mean}$, and $\overline{I} = I - I_{mean}$ and divide the correlation by the root sum of the squares of each window. This equation, shown below, has the operational effect of yielding a normalised cross-correlation matrix (NCC) which has values from 0 to 1.

$$CM = \frac{\sum_{i=-N}^{N} \sum_{j=-M}^{M} \overline{R}_{sq}(j,i) \overline{I}(x+j, y+i)}{\sqrt{\sum_{i=-N}^{N} \sum_{j=-M}^{M} (\overline{I}(x+j, y+i))^2} \sqrt{\sum_{i=-N}^{N} \sum_{j=-M}^{M} (\overline{R}_{sq}(j,i))^2}} \tag{3.8}$$

The normalisation corrects for intensity differences in different regions of the image. A more efficient way to compute the NCC matrix is via Fourier domain methods. If I is a fixed image, and R is the moving image, and R_{rev} is the reversed moving image, then the NCC matrix (γ) can be efficiently computed using [14, 129, 130]:

$$\gamma = \mathcal{F}^{-1}(\mathcal{F}(I)\mathcal{F}(R_{rev})) = \mathcal{F}^{-1}(\mathcal{F}(I)\mathcal{F}^{*}(R)) \tag{3.9}$$

this is possible since, for a real signal, time-reversal of that signal is the complex conjugate of the Fourier transform [14]. Fourier domain operations like this avoid looping and are more amenable to real-time processing [14].

The location of the maximum of the NCC matrix corresponds to the shift. A

final note is on edge handling. The edges can either be neglected, where the sliding window will never be centred on an edge. Alternatively, the image can be cyclically repeated. They can also be zero-padded, which is used exclusively in this thesis for this case since we know there is no intensity outside the field of view.

A visual example of the NCC method is shown in Fig. 3.19. NCC has superior operational properties that make it useful for fibre bundle endomicroscopy, as well as a few drawbacks. First, it is intensity-invariant, and when a probe moves, features which were in the darker edge of a bundle can be matched to their new location in the brighter centre. Second, the location of the maximum of the NCC matrix relative to the location of the reference window gives the shift. As far as disadvantages, it is not sensitive to rotation and has no inherent way to handle tissue deformation such as sheering or artefacts moving against the direction of the probe.

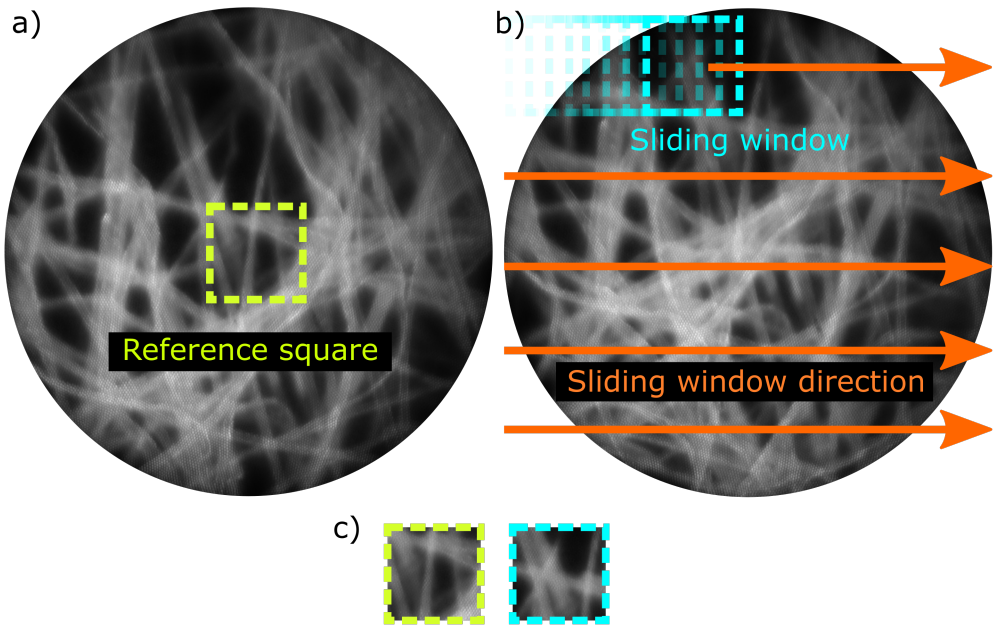


FIGURE 3.19: Sliding window operation. (a) reference image, (b) comparison image, (c) reference square and window extracted from comparison image. A reference square is extracted from an image, and a sliding window "scans" across the second template image. At each translation of the sliding window, the correlation is computed. NCC is intensity normalised to correct for global intensity variation.

An easy way to show correlation, matching, and NCC (Fig. 3.20) is to simulate a linear camera which has taken two snapshot images of a feature moving along the line. The two images have an intensity which is brighter near the centre, and dimmer near the edges, this is meant to simulate the same behaviour in a fibre bundle. The intensity dropoff is determined by multiplying a Gaussian with a specified pixel width against the images. The filter window selects the feature and then computes the correlation, matching, and NCC matrices which are plotted with the same dimensions. With the example shown, any advantage presented by NCC is not obvious, the contrast in the correlation and matching matrices is much higher than the NCC, and matching has a resolved local minimum. If anything, NCC appears to be the inferior method. The advantage of NCC lies in its scale-invariant properties, become apparent when the intensity profile across the images are much greater (Fig. 3.21), this simulates scattering in thick tissue. In this example, matching has completely broken down, although correlation and NCC matrices approximately maintain the resolvability of the peak. This simulation is not perfect; the filter window is chosen to be the feature, whereas, in reality, the entire field would contain features and background and would be sampled at random. What it does show is the breakdown when detecting the same feature, which is attenuated or amplified during sampling.

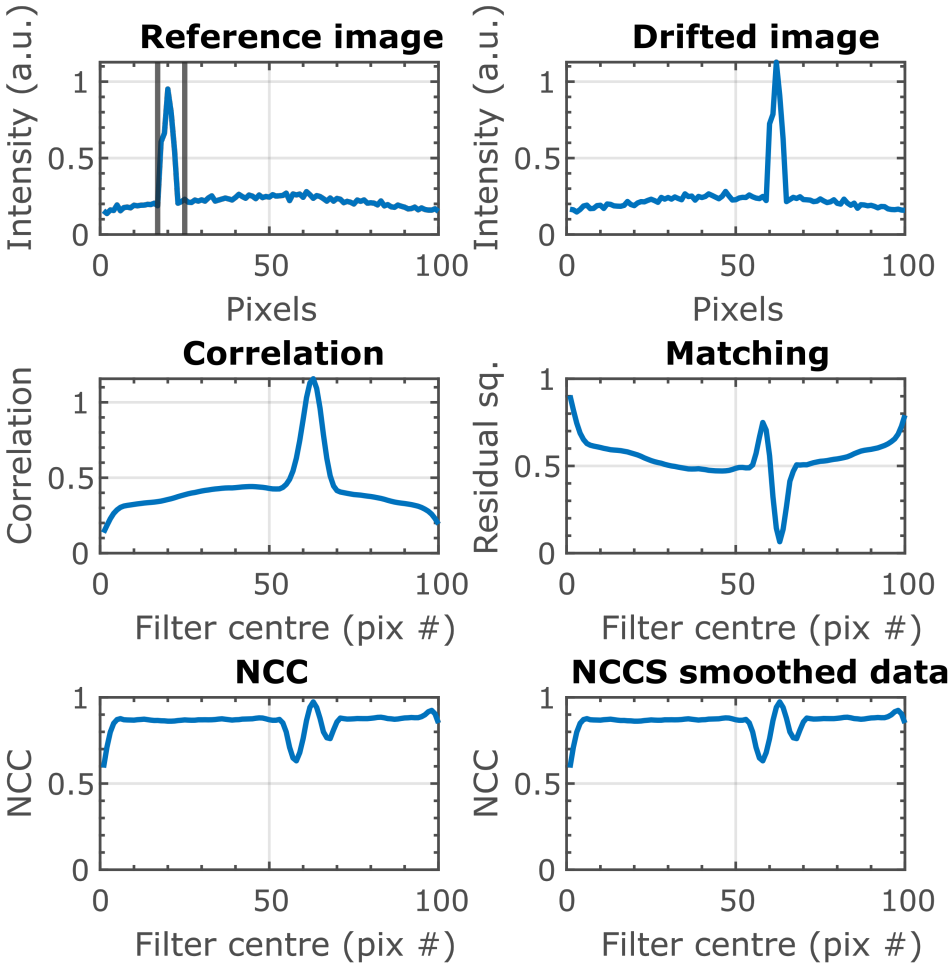


FIGURE 3.20: Comparison of correlation, matching, and NCC approach. Intensity dropoff gaussian width = 50 pixels. Simulation of two images of a moving feature acquired with a linear CCD. The intensity drop-off near the edges is simulated by a multiplication by a centred Gaussian.

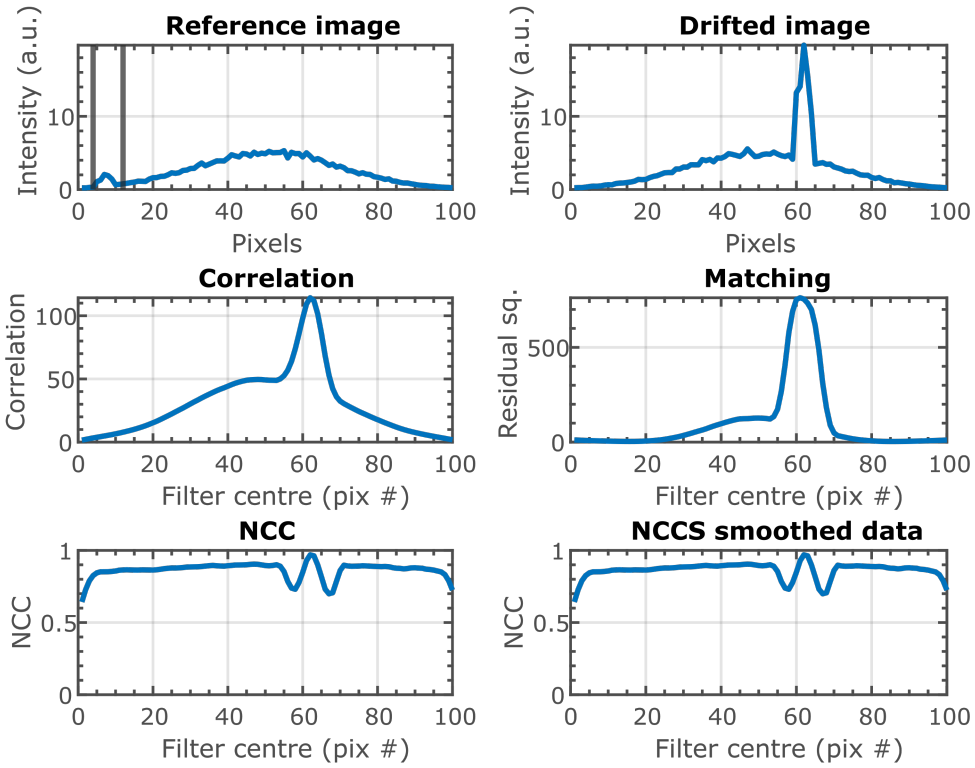


FIGURE 3.21: Comparison of correlation, matching, and NCC approach. Intensity dropoff gaussian width = 20 pixels. Simulation of two images of a moving feature acquired with a linear CCD. The intensity drop-off near the edges is simulated by a multiplication by a centred Gaussian.

3.4.2 Limitations

Registration errors can lead to mosaics that do not faithfully resemble the ground truth. Once a registration error is spotted, then the mosaics need to be cleared. Two ways to identify registration errors are a simple intensity check to identify a probe not in contact with a sample, and a threshold value checks for the maximum of the NCC matrix. The rigid registration approach leads to compounding errors due to the interframe error in each placement compounds. A stream can be later processed using more complex offline methods to compensate for this shift with global registration correction, and tissue deformation.

3.5 Summary

Cladding suppression presented five unique methods of removing the honeycomb structure in the images. The fastest was the Gaussian low pass filter, spatial frequency domain filtering with a disk filter both having the same number of processes and neither requiring looping across images. The best resolution and contrast maintaining was core localisation and interpolation. In terms of simplicity of design from the most simple to least, the methods were: Gaussian low pass, median filter, spatial frequency disk, local area comparison, localisation and interpolation. Arguably, the best method for real-time processing is the spatial frequency domain filtering with a disk filter. Disk filtering exhibits highly selective removal of cladding in Fourier space. The best method for precision is localisation and interpolation. A good "all-around" process which delivers good speed, performance, and is simple to implement is local area comparison.

Mosaicking can extend the useful area beyond that of the probe area alone. Real-time mosaicking using matching via NCC is robust to intensity variation across a probe, and dead-leaf image insertion allows for real time image insertion. Real-time mosaicking is limited by cumulative registration error. More robust techniques can then be used offline to correct for global registration and distortion, improve blending, and in some cases improve resolution.

Chapter 4

DMD based structured illumination endomicroscopy

4.1 Introduction

This chapter presents the development of a structured illumination (SIM) endoscopic microscope using a digital micromirror device (DMD), including theoretical and numerical modelling and, for all models, experimental verification. The basic idea of SIM is to suppress out-of-focus light through the spatial modulation of excitation light intensity at the proximal face of the bundle. This chapter begins with a brief technical review of three-phase demodulation, a common way of obtaining optically sectioned images using SIM.

The background section focuses only on the use of structured illumination for optical sectioning, and not for improvements in lateral resolution. The reason for this is when adapted to fibre bundle endomicroscopy lateral resolution is limited by the spacing between the cores of the bundle. In systems without a distal lens lateral resolution is typically an order of magnitude greater than the diffraction limit so improvements in lateral resolution would not be realised.

An optical system schematic and a technical procedure to implement SIM using a DMD are then presented, as are numerical simulations and experimental

characterisations of a DMD endomicroscope. The simulations are performed using two methods, a comprehensive, home-built core-wise intensity propagation approach using the Fraunhofer approximation and a computationally efficient Gaussian convolution. Key parameters which are determined are the bundle contrast transfer as a function of modulation frequency, the edge response function (ERF), and its derivative, the line spread function (LSF), the OTF, and the axial response with defocus.

The system is validated, showing clear background rejection in images of highlighter stained lens paper with a fluorescent background. Performance with animal tissue is also shown using *ex-vivo* porcine stomach, which was fluorescently stained with 50 g/ml acriflavine hydrochloride. It is further demonstrated that when the modulation depth and OTF are known, the system is capable of depth resolving. Three depth sections of a three-plane image are demonstrated. Further work on reducing motion artefacts from this system is presented in the next chapter.

4.2 Background

4.2.1 SIM with three-phase demodulation

Structured illumination microscopy reported by Neil, et al. [131] is a modulation and demodulation scheme enabling imaging with the complete removal of grid artefacts. Early work such as that pioneered by Schmidt, and Neumann [132, 133]

used carefully aligned gratings. Basically it providing sectioning when a beating pattern was superimposed on the in-focus portion of resulting images. Frequency domain processing then allowed for the isolation of the in-focus features, however grid artefacts that resulted from this beat pattern was not completely removed. It is discussed later (Sect. 4.6.2), that although SIM achieves complete removal of residual modulation, this is only true when using sine waves instead of square waves.

SIM is a three frame acquisition process, and each raw frame is acquired with a one-dimensional modulation pattern. We can express the modulation, $s_i(x, y)$, as

$$s_i(x, y) = \frac{1}{2}[1 + m\ \sin(\nu x + \phi_i)] \tag{4.1}$$

where m is the modulation depth, which ranges from 0 (no modulation) to 1 (complete modulation), ν is the spatial frequency of modulation, chosen based on desired optical-sectioning strength, and $\phi_i = \frac{(i-1)2\pi}{3}$, i is the image number (1,2,3) in the three frame sequence [131]. The $1/2$ scaling, as well as the first term, is to ensure there are no negative intensity values.

Each subsequent image after the first is shifted by $\frac{1}{3}$ of a period. Once the three raw images, I_1, I_2, I_3 have been acquired, the modulation is removed, and an optically sectioned image I_{SIM} obtained, via three-phase demodulation[131, 134–136]:

$$I_{SIM} = \sqrt{(I_1 - I_2)^2 + (I_1 - I_3)^2 + (I_2 - I_3)^2} \tag{4.2}$$

A complete mathematical treatment is available in Ref. [131]. To show the in-focus case with a modulated object, simple substitution of $I_{object}\sin(\nu x + \frac{(i-1)2\pi}{3})$, where I_{object}, is the intensity of the unmodulated object, into Eqn. 4.2 yields:

$$I_{sim} = I_{object}\frac{\sqrt{2}}{3}\sqrt{\begin{aligned}(sin(\nu x) - sin(\nu x + \frac{2\pi}{3}))^2 + (sin(\nu x) - sin(\nu x + \frac{4\pi}{3}))^2 + \\ (sin(\nu x + \frac{2\pi}{3}) - sin(\nu x + \frac{4\pi}{3}))^2\end{aligned}} \tag{4.3}$$

and after reduction:

$$I_{sim} = \frac{\sqrt{2}}{3}I_{object} \tag{4.4}$$

If an image has in (I_{in}) and out-of-focus (I_{out}) components, each modulated raw image is given by:

$$I_{raw} = \frac{1}{2}[1 + m\ \sin(\nu x + \phi_i)]I_{in} + I_{out} \tag{4.5}$$

where the out-of-focus light is the same in all three images and is subtracted from each difference in Eqn. 4.2, reconstructing only the in-focus contributions.

The modulation formula (Eqn. 4.1 has an added term of $\frac{1}{2}$. The shift can be thought of as a DC term and is subtracted inherently in Eqn. 4.2.

What follows from the sinusoidal terms is that the modulation depth attenuates faster with a higher ν. Defocus can be approximated by blurring an image with a Gaussian kernel [5]. Moreover, once the transfer properties of an optical system are known, the Gaussian kernel which corresponds to each depth is known. The sine waves can then be directly convolved with that kernel to get the effective modulation depth; this also gives the axial response to defocus. The

axial response is the intensity response, after reconstruction, of a system to a thin, infinitely long and wide sheet moving away from focus. It is normalised by the in-focus intensity. While it may seem redundant to consider the OTF and axial response separately, the bundle samples incident light by averaging over each core, this behaviour can only be accounted for by taking the three images, which accounts for the three-phase shifts. For the optical system used in this chapter, the width of the Gaussian kernel needed at each depth is simulated and experimentally verified, and mapped to depth in Fig. 4.17.

4.2.2 SIM endomicroscopy

The first demonstration of a SIM endomicroscope was by Bozinovic, N. et al. (2008) [43]. A multi-core fibre bundle was placed at the focal plane of a benchtop SIM microscope. The schematic is shown in Fig. 4.1. The modulation pattern is generated by reflecting laser light off of a spatial light modulator (SLM) which is relayed to the proximal face of the bundle. Grid patterns can also be generated with other methods such as two-beam interference [137]. The returning fluorescence is then separated by a dichroic beam splitter and emission filter and imaged onto a camera. The pattern refresh rate was 10 Hz, with latency included the imaging rate was 2 frames/s.

Sample images have differences when imaged in free space vs through a bundle, and this can be seen in the Fourier domain. One difference is a ring corresponding to the cladding, and the second is a dampening of frequencies in the OTF near the Nyquist cutoff of the bundle. A Fourier characterisation of the endomicroscope used in this chapter, which identifies those features, is shown in

Fig. 4.9.

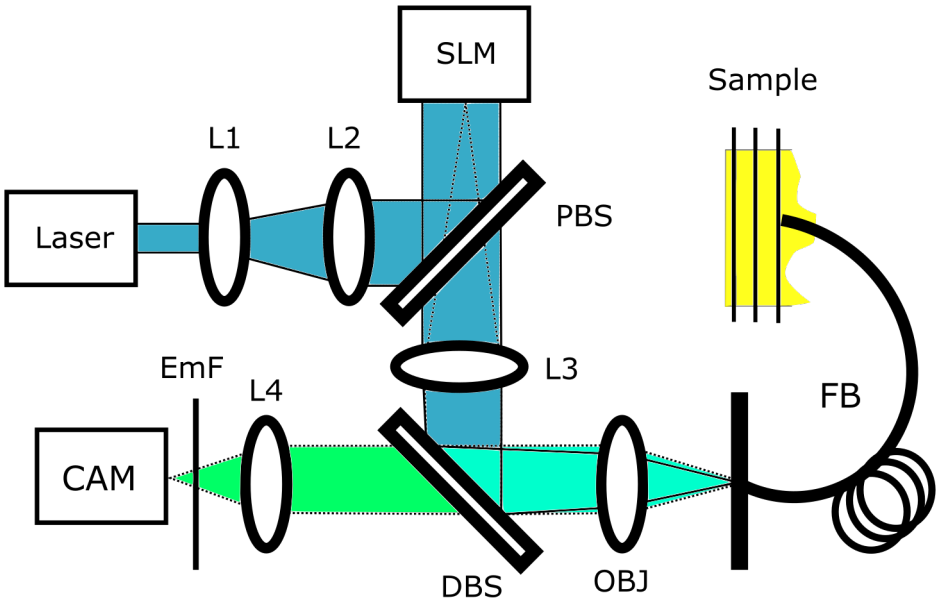


FIGURE 4.1: SIM endomicroscope. Laser light is reflected off of an SLM and onto the proximal face of a fibre bundle via a beam expander (L1, L2), polarising beam splitter (PBS), L3, DBS, and objective. The returning fluorescence emission is then imaged onto a camera via the objective and a tube lens. Based on the design described by Bozinovic, N. et al. in Ref. [43]

During imaging Bozinovic, N. et al. noted residual artefacts. These have been well described in SIM microscopy [138, 139]. They come from a number of sources, such as inconsistent illumination when the modulation depth of the sine waves is different for each of the three raw patterns, differences in phase angles, and trapezoidal-shaped fringe patterns. One solution reported to reduce artefacts was to introduce high pass filtering. Modulating a raw image means that in-focus features are shifted to the sidebands, artefacts can be reduced by taking a Fourier space representation and high pass filtering features just slightly above the 0th frequency. Artefacts are explored in more depth in the limitations section (Sect. 4.6), and a different approach is proposed.

Further development in SIM endomicroscopy was presented by Kyrish, M. et al. (2012) [63]. The endomicroscope used a motorised grating to generate line

pairs. Unlike an SLM, a fixed grating was used, meaning only a single spatial frequency was accessible, without changing the grating. Similar reports use a grid attached to a piezoelectric device in transmission in a conjugate plane to the sample [140]. Furthermore, if a grating is translated at a constant speed, intraframe motion artefacts can result. The lensed system, with a custom-built plastic achromatic objective, used incoherent LED illumination along with a proflavine contrast agent to *ex-vivo* image murine mammary tissues.

It is useful to select modulation frequency to balance the signal and optical sectioning strength. Higher modulation frequencies are desirable for improved sectioning; however, they are often diminished by the OTF and sampling by the bundle [42, 136], the subsequent reduction in contrast, reduces the dynamic range. In 2015 Keahey, P. et al. presented a configuration based on a motorised grating which was capable of three different line pair widths. Measured at the distal plane these were 100, 50, 25 μm [42]. They reported an image acquisition rate of 11 frames/s with an effective post-reconstruction rate of 3.7 frames/s. Keahey also reported diminished signal levels depending on grating line spacing, gratings having respective in-focus signal intensities of 90, 68, 33%. Although the signal at 100 μm was 90%, the HWHM of the axial response was a factor of 2 broader between 100 and 25 μm. Furthermore, motion artefacts were also reported when the probe shifted between acquisitions.

Motion artefacts which occur between acquisitions can be reduced using higher speed acquisition [42]. One way to increase the speed, which has been developed for benchtop applications, but not previously to endomicroscopy is by using a

digital micromirror device for fast SIM [141]. A DMD, basically an array of high-speed switchable mirrors, is described in depth in Sect. 4.3.2.1. The reported acquisition rate for a conventional SIM microscope was 50 frames/s, leading to an imaging rate of 16.7 frames/s [141].

Keahey proposed using a DMD for structured illumination endomicroscopy to improve speed, and consequently reduce motion artefacts, and for the ability to tune modulation frequency on-the-fly. DMDs do have drawbacks as were reported by York, A. et al. [142] when using one for super-resolution structured illumination. The first is that DMDs are not zero-order diffraction, meaning the use of a coherent light source will reduce contrast at the focal plane, and the relative effects of this increase as the modulation frequency increases. This is mentioned for generality and is not a factor in LED based approaches like the one presented in this chapter. The second drawback is that DMDs have a finite contrast ratio, and even with all pixels in the off state, light still reaches the sample. These limitations were reported to potentially limit the compatibility of a DMD based endomicroscopy with staining agents [141].

In 2015 Ventalon, C. et al. [44] in a conference abstract, reported a different application of structured illumination, to depth section images using a lensless bundle. Using several different modulation frequencies, the author showed three depth sections, 0-15, 15-30, and 20-50 μm. This was a subtractive approach which involved computing the difference in images with well-chosen grid modulation. As a contact mode bundle, with no focusing power, planes near the bundle had a higher resolution than planes deeper in the sample.

This thesis now reports, characterises, and validates the first DMD configuration of a SIM endomicroscope. This system allows for tunable spatial frequencies, does not lead to artefacts caused by phase misstepping, and is capable of speeds of 40 processed frames/s. Image performance is shown for tissue paper separated from a fluorescent background by a plastic sheet, and bovine stomach stained with acriflavine hydrochloride. Depth sectioned slices, as part of a subtraction based configuration, are also presented.

4.3 Method

4.3.1 System description

A schematic is shown in Fig. 4.2. Light from the blue LED of a DMDs, (Fig. 4.3.2.1) built-in illumination system is directed through a 20x/0.25 NA finite-conjugate objective (Newport), a 20 mm collimating lens (Thorlabs ACL2520U), an excitation filter (Thorlabs FESH0450), and reflected off of a 490 nm cut-on dichroic mirror (Thorlabs DMLP490) onto the back aperture of a 4x / 0.25 NA finite-conjugate objective. It is then transferred through a 1 meter long fibre optic bundle (Sumitomo IGN-08/30) in pixelated form. The returning fluorescence emission is imaged onto a camera (Point Grey Flea3) via the 4x objective, dichroic beam splitter, and an emission filter (Thorlabs FELH0500). A microcontroller (Pyboard 1.0) is used to control triggering of the camera and DMD (Fig. 4.5). A data acquisition (DAQ) card (NI USB 6008) provides communication between the controlling PC and the micro-controller. At the object plane, the width of each camera pixel corresponds to 0.34 μm, with the 4 μm inter-core spacing of the bundle this provides

better than Nyquist sampling. The width of each DMD pixel corresponds to 8.6 camera pixels (2.9 μm) measured at the object plane. The inter-core spacing (ICS) of each core was $\sim$ 3.2 μm. Images are acquired in contact mode only. An image of the lines projected on the proximal face of the bundle is shown in Fig. 4.4.

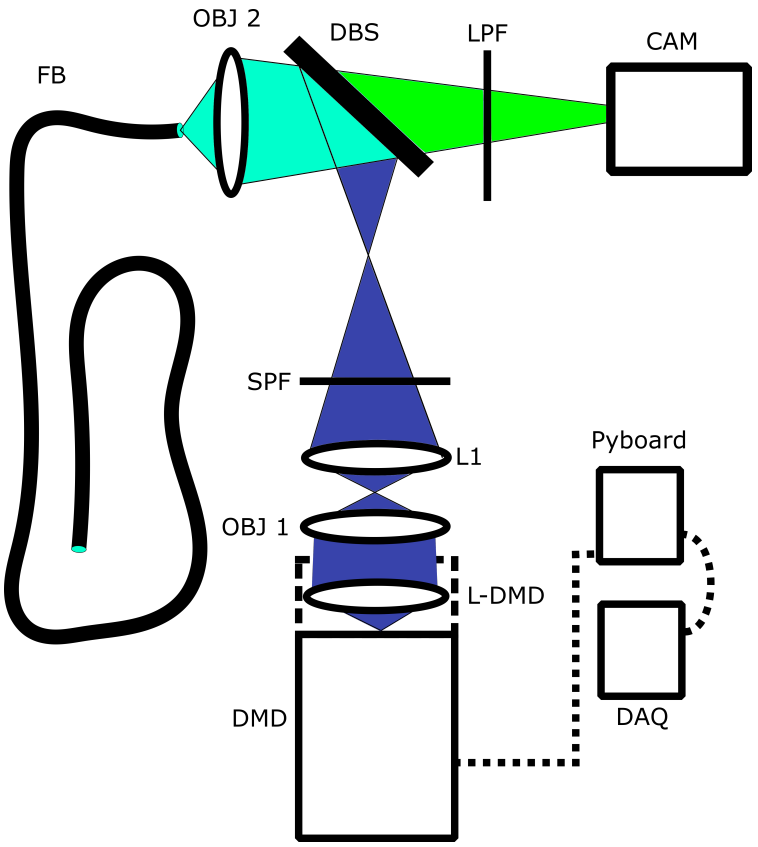


FIGURE 4.2: Schematic of endomicroscope. DMD - Digital micro-mirror device, OBJ 1 - 20x, OBJ 2 - 4x objective, L1 - 20 mm collimating lens, L-DMD - lens built into DMD, SPF - Short pass Filter, LPF - Long pass filter, DBS - Dichroic beam splitter, FB - fibre bundle.

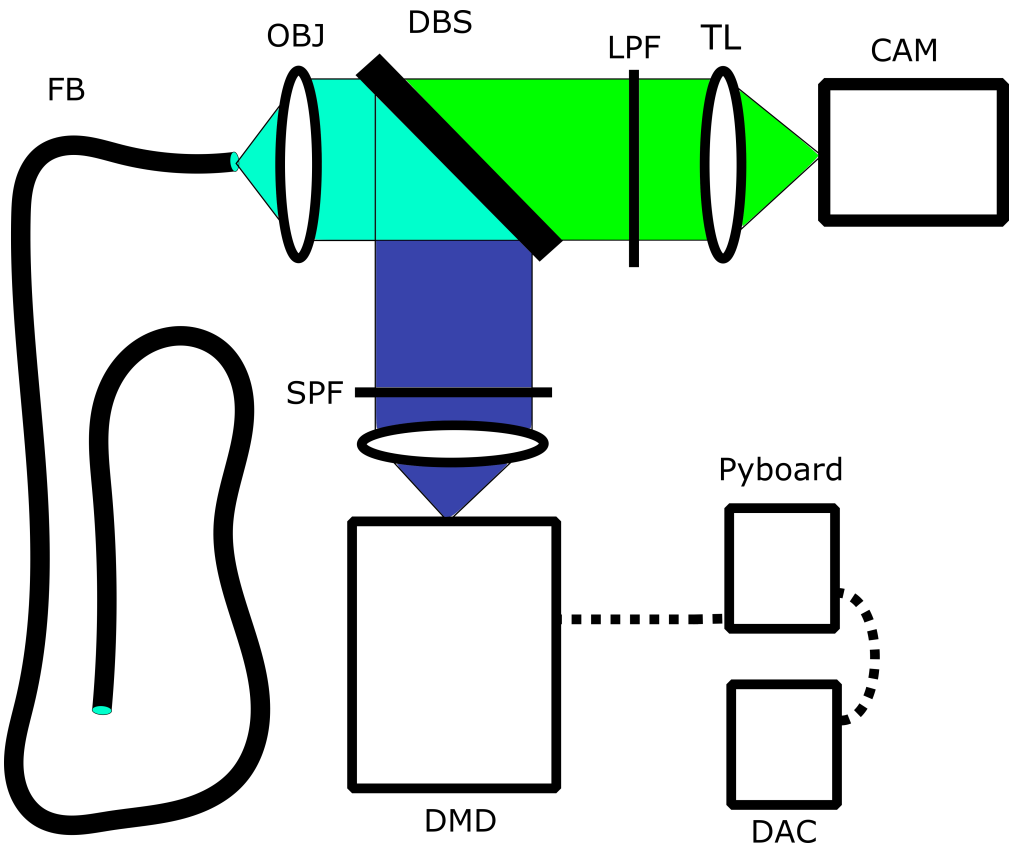


FIGURE 4.3: SIM System without a built in DMD lens. DMD - Digital micro-mirror device, OBJ - infinity corrected objective, SPF - Short pass Filter, LPF - Long pass filter, DBS - Dichroic beam splitter, TL- tube lens, FB - fibre bundle.

Additional lenses are required near the DMD in Fig. 4.2 to compensate for the DMDs built-in lens. The system can be made simpler by removing the lens element, a schematic showing this proposed, but not implemented, configuration is shown in Fig. 4.3. There are a few benefits to this configuration: the use of infinity space for the filters to prevent changes in magnification if a filter change is desired, as well as fewer optical elements and no aberrations introduced by the DMDs lens (address further in Sect. 4.3.2.2).

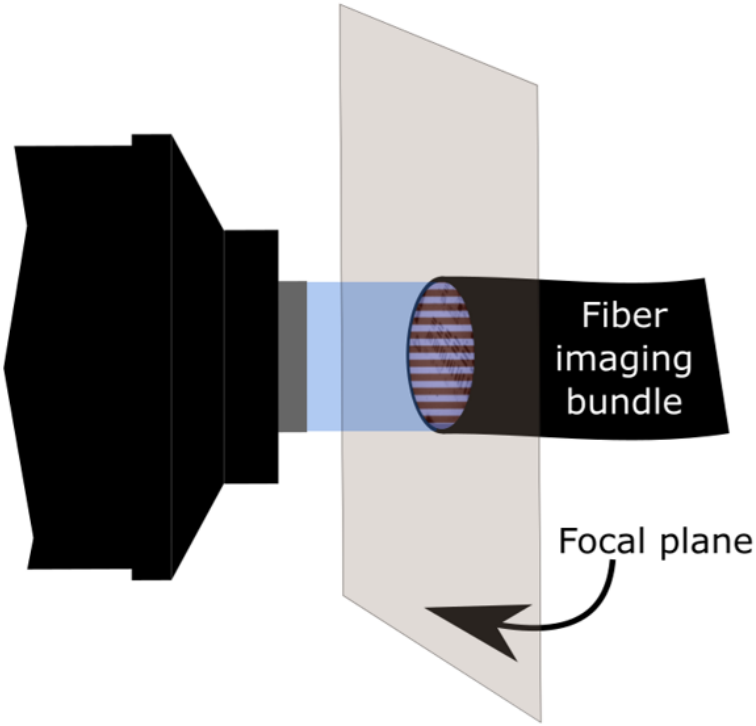


FIGURE 4.4: Lines projected onto the proximal face of the fibre bundle.

4.3.2 Generating line patterns

This section will discuss some of the practical challenges in implementing SIM on a DMD, including how we can optimise the lines to the rotated square-pixel structure of the DMD array.

4.3.2.1 Digital micromirror device

A DMD (TI DLP 3000) is shown in Fig. 4.5. The device is an electronically controllable array of mirrors. The square pixels have dimensions of 7.6 μm, and they are fabricated on a diamond array. Light is incident from the side. Binary amplitude modulation is done by turning pixels on or off, and intermediate grey values can

be generated by dithering the mirrors for a set amount of time, i.e. 1/8 of an exposure period.

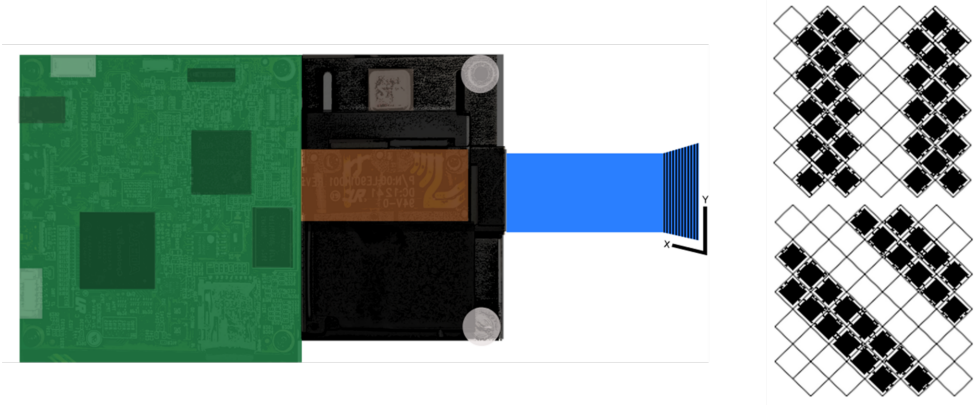


FIGURE 4.5: Image showing a digital micromirror device as well as the pixel layout. The most narrow lines attainable (optimal sectioning) requires the lines to be oriented with the DMDs rectangular array.

Patterns can be projected by the DMD in a few ways, first, and the primary mechanism used in this thesis, is via onboard pattern control. Patterns are loaded on over a USB 3 data interface. To display patterns, the DMD then loads the pattern from memory and updates the mirror positions. The LED is always on when pattern mode is active. The number of patterns that can be loaded onto the DMD is determined by the bit depth of the patterns. 1-bit patterns support 96 onboard, while 8-bit patterns support 12 onboard. The patterns can either be triggered using a camera, the DMD can output trigger signals, or external triggering can be used. After each trigger pulse is received, 365 µs is required for the next pattern to be displayed (230 µs to load the bit plane, and 135 µs buffer rotate overhead). Patterns can also be loaded using an HDMI interface, although precise bit-plane control is required to change patterns at a high frame rate. Precise bit-plane control isn't available when using a DMD as a second monitor with a Windows 10 based operating system, and Intel I5 series graphics. A device that allows for individual bit plane control, is required.

4.3.2.2 Discretisation error

Aside from line defocus artefacts, discretisation error occurs when lines, generated by computer, are down-sampled to the pixel structure of the DMD. Discretisation error results from a few sources, firstly, when generating lines on a discrete grid, the number of pixels, must have 3 and 2 as common factors, secondly the patterns must not be rotated on this grid. As will be discussed in Sect. **??** neither criteria can be completely satisfied with a freehand moving probe.

Neglecting the effects of the bundle, it is desirable to have lines generated on the DMD as narrow as possible since optimal SIM performance is achieved with the highest available frequency. Lines must have a width, in pixels, with 2 and 3 as common factors. Two must be a common factor since there needs to be equal on and off areas, and three since the spatial period of the lines must be phase advanced 1/3. This means the minimum number of pixels for a SIM system needed is 6. Fig. 4.6 shows the pixel representation of a 6-pixel system; the theory is not defined using line pair widths, in pixels, that do not have 2 and 3 as factors.

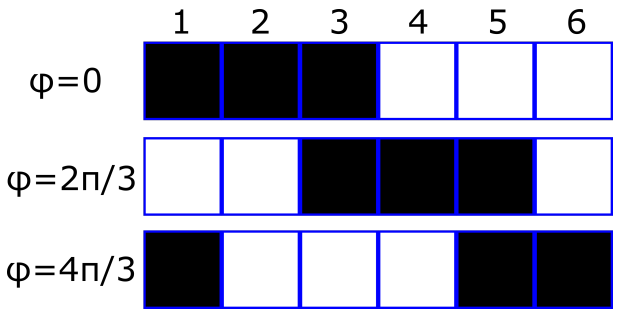


FIGURE 4.6: Pixel representation of a single DMD row for the three-phase SIM sequence.

Induced sampling error results when lines are not perfectly oriented to the

grid array. A common design for digital micromirror devices is to have a diamond pixel structure[143]. In contrast, arrays used to generate lines on a pc will have a rectangular structure. Discretisation artefacts, which result from sampling a rectangular array with a triangular, are avoidable when attention is paid to the DMDs mapping scheme. In the motion compensation chapter (Chap. 5), we rotate the patterns to orient them with the movement of the probe—the artefacts that result come from residual line artefacts, and difficulties introduced by discretisation. Discretisation error is caused since the lines can be perfectly oriented to the diamond structure of the DMD array.

A few key observations can help minimise artefacts: first, residual line artefacts result when square waves blur and modulation appears trapezoidal [138]. This occurs when using line modulation, which when decomposed into a Fourier sine series has harmonics truncated by the OTF. Sect. 4.6 describes that, when the optical transfer function only passes the first two sine Fourier components, the fundamental frequency and 3x the fundamental, then there are no line artefacts. Therefore blur can be deliberately introduced to generate sine waves at the proximal face of the bundle. Aberrations, however, introduced by the DMDs built-in lens may limit the effectiveness of this in the system used in this chapter, but pose no downside generally. The second observation is that when using rotated patterns the artefacts that result from discretisation error, including the problem with non-binary values, occur at the edges and the simplest way to reduce this artefact is fewer edges, or lines with a larger width, the trade-off, of course, being sectioning strength.

In Sect. 2.5, several previously reported limitations of a DMD were discussed.

Including that using an incoherent light source will reduce contrast at the focal plane, and that even with mirrors off some still light transmits. The DMD modulation is now characterised, taken at the focal plane. The discretisation rules that apply to DMD grids, 2 and 3 must be common factors of pixel width, also apply to bundles. By showing near-complete modulation when LPW $<<$ 6 ICS, it can be argued that DMD effects alone do not greatly diminish contrast. Lines are generated at 16 μm ($\sim$ 3 ICS), and the image is then divided by a widefield image, to get the modulation depth. The process of normalisation is detailed in Sect. 4.5.1. Fig. 4.7 shows a clear minimum at 0 and maximum at 1, and sharp intermediate peaks. This indicates complete modulation ($m = 1$) at the focal plane indicating that DMD effects, in this case, do not adversely contribute to diminished modulation depth.

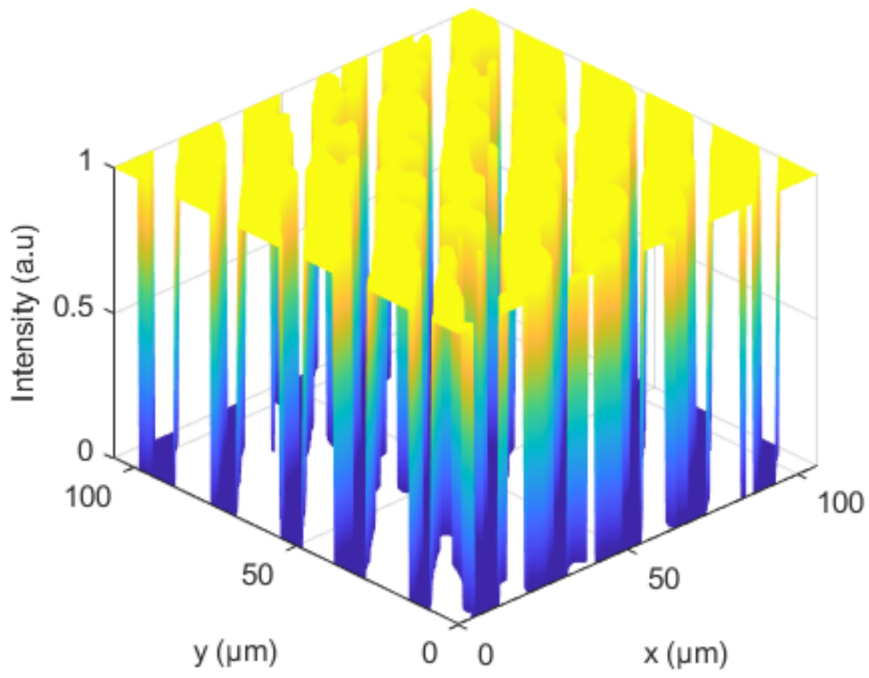


FIGURE 4.7: Fluorescently labeled smooth plastic cap acquired without a bundle. Intensity normalised by a widefield image. Illuminated with a 10x infinite conjugate objective.

4.3.3 Line modulation scaling factor

Line patterns have benefits when implementing with a DMD. Namely, they can be transmitted as binary 1s and 0s since the mirrors are either on or off. Sinusoidal modulation, with narrow exceptions at the peaks and troughs, has intensity values that vary between 0 and 1. It is possible to generate sinusoidal patterns, by dithering the mirrors, where the DMD turns mirrors on for a short time then off. The precision of patterns means they require a higher bit depth, consuming more memory onboard the DMD. It isn't necessarily critical to use sinusoidal modulation. The simulation in Fig. 4.8 demonstrates what has been well reported in the literature, for in-focus reconstructions, square waves are equivalent to sine waves for a perfectly aligned grid. The figures show an intensity profile from a single frame of the reconstruction, as well as the reconstruction of an in-focus image.

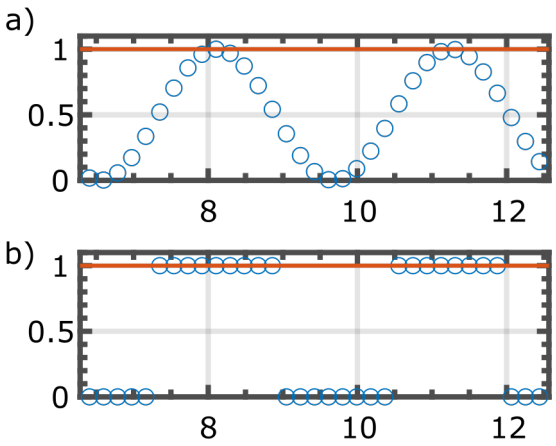


FIGURE 4.8: Comparison in (a) sine wave and (b) square wave SIM reconstructions. Modulated in-focus plane. Solid line - intensity profile from reconstructed image. Dotted line - intensity profile from a single frame. Both methods accurately reconstruct the plane. Scaling factors: $\frac{2\sqrt{2}}{3}$ (Sine wave), $\frac{1}{\sqrt{2}}$ (square wave)

The formula for three phase demodulation is given by [134]:

$$I_{sim} = S_c \sqrt{(I_1 - I_2)^2 + (I_1 - I_3)^2 + (I_2 - I_3)^2}. \tag{4.6}$$

Hagen, N., et al., discusses the variety of coefficients, s_c, in the literature including 6 papers which use $\frac{\sqrt{2}}{3}$, appropriate when the modulation depth, m=1, and the sinusoidal modulation in Eqn. 4.1 is not scaled by $\frac{1}{2}$. Hagen also refers to other authors which use $\frac{1}{3\sqrt{2}}$, $\frac{1}{\sqrt{2}}$ or $3\sqrt{2}$ without stating a reason. By simple substitution of $I_i(x,y)$ with $i = 0,1,2$, it is possible to show that $\frac{2\sqrt{2}}{3}$ accurately reconstructs f(x,y), but this is only true for the case of sinusoidal illumination.

Although grid patterns are commonly used, a coefficient for the quantitative sectioning formula is not often explicitly reported for line illumination. That said, it has been used before without explanation [135, 144]. The coefficient can be determined simply by considering a single pixel, across the three images, when imaging an in-focus plane of unit intensity. The single-pixel will either have some permutation of [0,1,1] or [0,0,1]. Substituting in any one of the two possible yields reconstructions of $\sqrt{(2m)}$, meaning $S_c = 1/\sqrt{2m}$ is the coefficient needed for quantitative reconstruction.

4.4 System characterisation

Similar to Bozinovic, N. et al., a Fourier space characterisation of a single SIM raw frame is presented in Fig. 4.9. There are clear regions of interest identified, such as the cladding ring and the Nyquist cutoff. The bright centre is the term corresponding to the dc term (since the sine waves are shifted), and the strongest sidebands can be seen more centrally, nearer to the location of the fundamental frequency, the damped harmonics are the result of using square instead of the sine wave modulation.

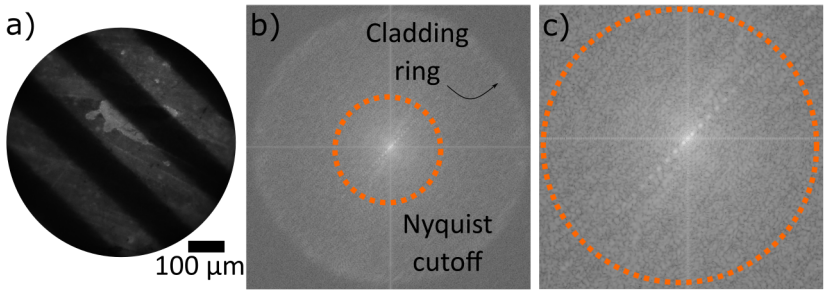


FIGURE 4.9: 300 µm spatial frequency grid modulation, measured at object plane, transmitted through 620 µm bundle (Fujikura FIGH-30-650S). (a) Raw image, (b) magnitude spectrum of Fourier transform, (c) enlarged region to show OTF cutoff.

4.4.1 Simulation description

Light transmitting through a bundle is different than in free space for a number of reasons, namely: the pixelation effects of the bundle, the NA of the fibre bundle, and the fraction of light transmitted by each core depends on impurities in the core, the structure of the core, and the angle light was incident at. The remainder of this chapter compares simulated results with experimental results. The simulations were performed by first controlling the modulation pattern projected on the proximal bundle face. The proximal pattern is then averaged over on a core-wise basis. Then, if applicable, transmission efficiency coefficients are applied to each core. Finally, the patterns on the distal face are generated in one of two ways, using either a home built simulation, described below or by using a Gaussian convolution with the PSF. Predictably they both yield similar results and are mapped via the LSF in Fig. 4.16.

4.4.1.1 Diffraction based simulation

The simulation works in the following way: x and y core locations of a bundle are generated. Then a high precision pattern (mesh elements >> inter-core spacing) is imaged onto the proximal face of the bundle. The pattern is then sampled by the bundle cores, with each core simply averaging intensity over its area. A distal image, representing the transmission through the bundle, is then generated. The light is then propagated to a mirror using diffraction integrals from a circular aperture at a known z distance, reflected back and re-sampled by the bundle. When considering how light propagates from the bundle, each core is treated as a uniformly sized circular aperture. The diffraction integrals are then solved using Fresnel and Fraunhofer approximations †. As an approximation for fluorescence imaging, which is incoherent, the field intensity, as opposed to the complex field, is added at the detection plane. The simulation does not include an NA cutoff.

The Fraunhofer and Fresnel approximations were used as solutions to the Fresnel diffraction integrals. These are far-field integrals, and the detection screen must be sufficiently far compared to the size of the aperture, and the detection screen must be a comparable size to the aperture[1].

This adaptable and modular code has benefits in terms of rigorousness. The code accurately replicates the intensity sampling behaviour of a bundle, losses through the cladding can be accounted for, the diffraction integrals can be generated separately balancing accuracy and performance, and there is no strict circular aperture condition. SIM patterns as well as x,y locations of cores are also

[1]** Bundle core locations generated using an API developed by M Hughes (https://research.kent.ac.uk/applied-optics/software-downloads/), all other code is self-built and included in the *SIM code* appendix.

embedded in a high precision mesh to minimise sampling error. In light of these benefits, it is very computationally expensive and requires $N_{cores} \times (n \times n)$ array addition operations, where N_{cores} is the number of cores, and n is the width, in pixels, of each diffraction integral.

4.4.2 Core shape and transmission efficiency

The sampling and transmission effects of a bundle are measured. Cores are characterised in terms of how (1) circular they are, and how uniform the (2) transmission effects are across each core of the bundle. Transmission effects are highly dependent on the angle of the incident light, and qualitative differences can be seen a probe is illuminated by a broadband white LED source in Fig. 4.10.

A high magnification image of a distally illuminated bundle was taken (Fig. 4.11(a)). To capture an image a bundle was fixed at the working distance of a 20x objective, a white light LED cell phone torch was used to illuminate the distal face, and an image was captured, using the same optical system as Fig. 4.2 with the objective replaced. A custom-built code is then used to segment the cores. The code worked in the following way: the median value of the distally illuminated input image was used as a threshold filter, then all objects less than 2 pixels (noise) were removed, the holes were filled in using MATLAB's imfill.m function, with holes selected as an option. This then made way for an image where cladding and core pixels were clearly isolated, and each shape was simply connected, without holes. This constraint allowed for the use of MATLAB's bwboundaries.m and returned an array of the boundaries of each core.

Next transmission efficiency for a white light source is reported. Each core was extracted from the image (Fig. 4.10), the cores area was then averaged over, to simulate bundle sampling, and a core-wise image was generated. The variation in the intensity of each averaged core indicates there is non-uniform transmission efficiency.

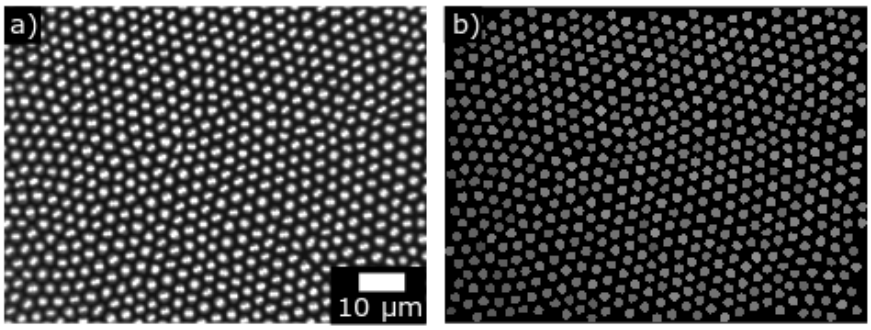


FIGURE 4.10: (a) Fibre bundle (Fujikura FIGH-30-650S) distally illuminated with a white LED. 20x objective. (b) Cores after threshold, and averaging. The difference in intensity represents core transmission efficiency.

To get a qualitative sense of the shape of each core, the *shape factor*, used as the figure of merit (FOM) for each core was computed:

$$SF = \frac{4\pi(area)}{perimeter^2} \tag{4.7}$$

Circles with a FOM of 1 are nearly perfect circles. The results of the image segmentation are shown in Fig. 4.11, and a histogram of FOM is shown in Fig 4.11(d).

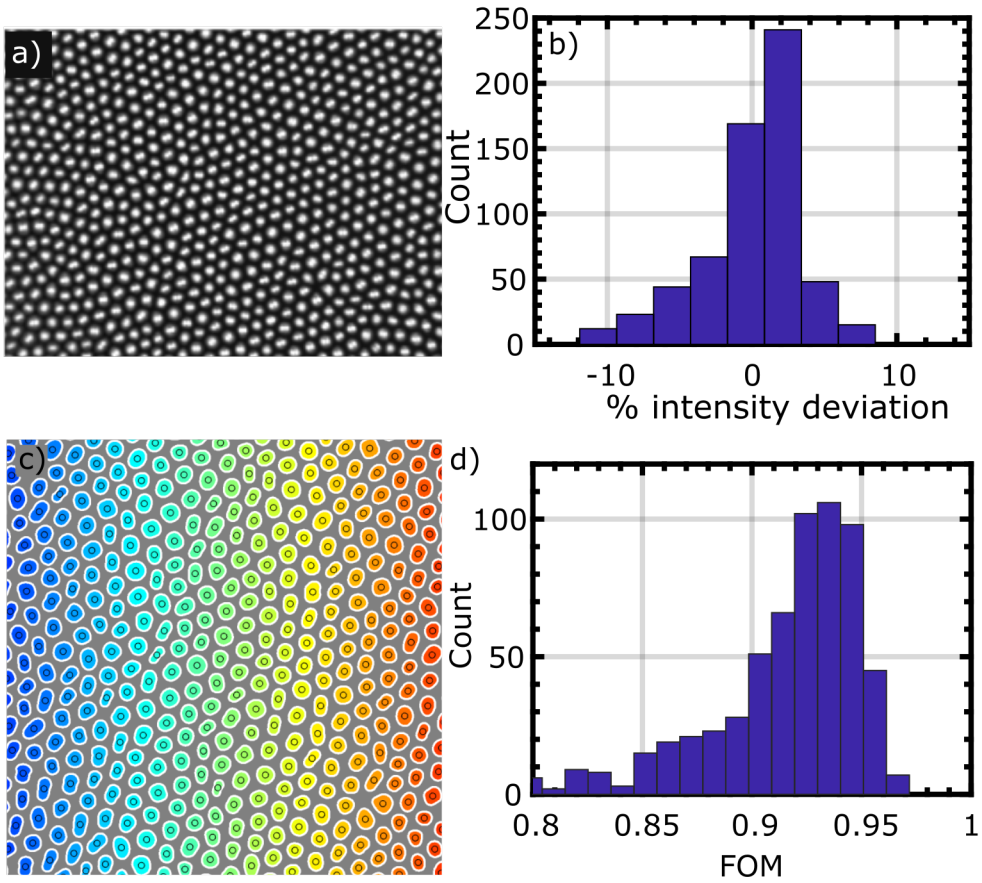


FIGURE 4.11: (a) 500 core image of a fibre bundle (Fujikura FIGH-30-650S) distally illuminated with a white LED, (b) histogram of % intensity deviation from mean, (c) segmentation results, (d) histogram of shape figure of merit (FOM); 1 is a perfect circle.

4.5 Experimental results and comparison with simulation

4.5.1 Bundle contrast transfer

Bundle sampling of high precision lines is not strictly equivalent to sampling periodically. Fibre bundles are formed from glass rods, which are stacked hexagonally, then turned on their side and drawn. During heating, the structure deforms, and global periodicity is broken. Since lines are formed on a rectangular

array, sampling by a pseudo-hexagonal grid means the lines can't be perfectly aligned with the hexagonal grid.

To guide optimal line pair selection, we first simulate, then experimentally verify, the reconstructed intensity after bundle sampling of in-focus lines of different widths. To understand why this will have an effect, consider the limiting case of an infinitely high and infinitely low spatial frequency. In the high spatial frequency limit, when sampled by the bundle, the distal pattern will be the same in all three images, and the reconstructed image will be just the noise. In the low spatial frequency limit, all the pixels in each of the three images will have exactly the same values, and while an image will be reconstructed, it will exhibit no added sectioning power.

The formula for raw images superimposed with the structured grating is:

$$I_{raw} = I_{object} \times (1 + m \sin(\omega x + \phi_i)) \tag{4.8}$$

where I_{object} is the intensity of the object, m, the modulation depth, ω, the modulation frequency, and ϕ_i, the phase term ($\phi_i = \frac{i2\pi}{3}$). To remove the contribution of the object, and correct for global intensity variations, each raw image can be divided by a widefield image (I_{object}). I_{object} must be a planar surface that is parallel to the bundle's face. The resulting image will have intensity centered around $I = 1/2$, and the global modulation depth, an average of local modulation depth, can be estimated as the difference between the minima and maxima. In cases where quantitative fluorescence information is needed, it can be useful to determine modulation depth at each pixel, and a numerical estimation approach for

this local modulation depth is developed in Ref. [134]. An alternative approach, and one that is used throughout this thesis, is to reconstruct SIM images from normalised images this effectively generates a 2D "modulation map". Image normalisation has the effect of removing I_{object}, leading to a map of local modulation depth estimates. The global modulation depth is estimated as the mean intensity of the reconstructed SIM image. The drawback is that modulation transfer effects are not considered alone, as the 2D "modulation map" considers the effect of bundle sampling effects as well (discussed further in Sect. 4.4.1.1).

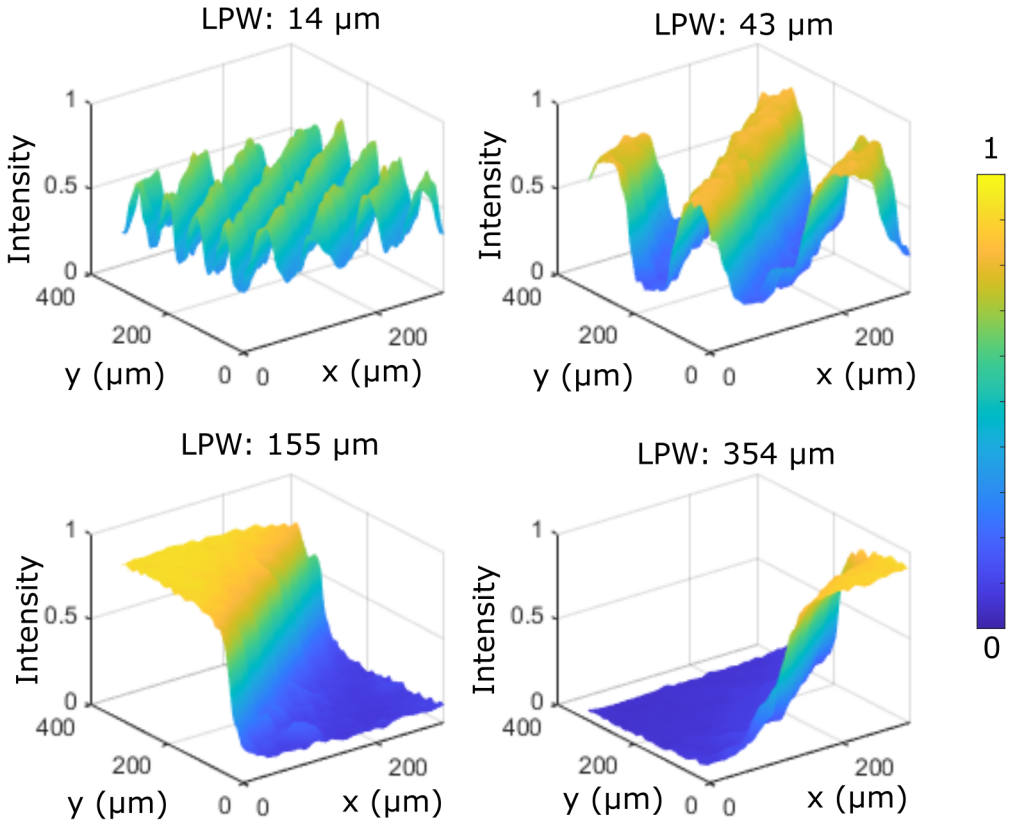


FIGURE 4.12: Modulation of a fluorescent metal plate in contact with a multicore fibre optic bundle (Fujikura FIGH-30-650S), using line pairs of different spatial frequencies. ICS of 3.2 µm. Cladding suppressed using a Gaussian kernel of width 1.5 µm. Raw images normalised by a widefield image.

The formula for intensity of each raw image is given by:

The modulation depth was experimentally confirmed. The distal tip of the fibre bundle was placed in contact with a smooth fluorescently labelled metal plate. A complete set of SIM images and a widefield reference (to account for intensity variation between acquisitions) were acquired. All four images are low pass filtered (0.5 × ICS) to remove the pixel structure.

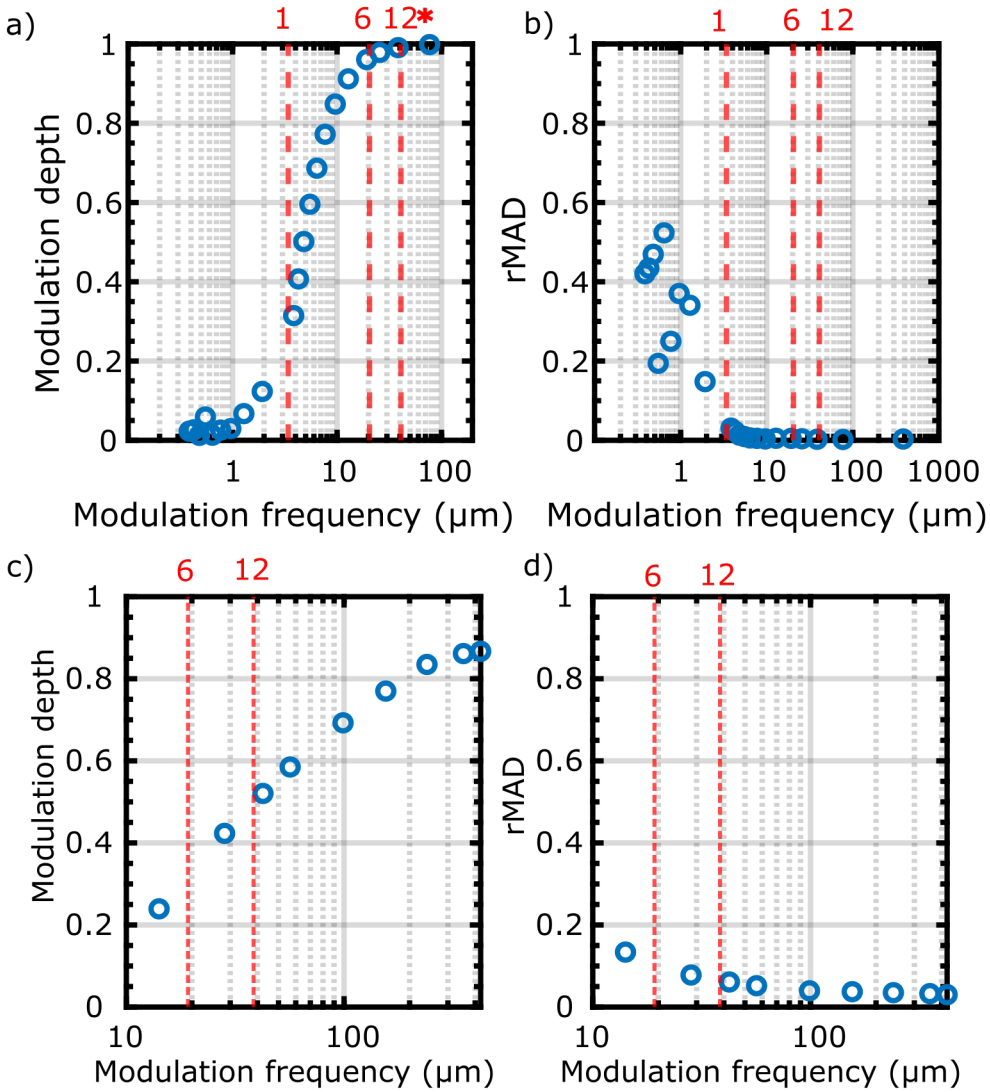


FIGURE 4.13: (a) Simulated reconstructed in-focus SIM intensity as a function of line pair width (b) and rMAD. Experimentally determined (c) reconstructed intensity, and (d) rMAD. Bundle with an intercore spacing of 3.4 μm. '*' and vertical lines indicates integer multiples of the core spacing.

The simulated modulation depth and relative error are shown in Fig. 4.13(a,b),.

What is clear from the simulation is that line pairs with a width 12 times the inter-core spacing is required to recover ideal reconstructions. In this simulation, noise and core transmission efficiency effects were neglected. The simulated intensity is $> 20\%$ in line pair widths below 2 $\times$ inter-core spacing, meaning the rMAD, being a relative measurement, is much higher. As the intensity increases, the relative error contributions fall off. The experimental results did not confirm the simulated results (Fig. 4.13(c,d)). Much greater line pair widths were required to recover contrast. What's more, modulation depth converged at around 95%. This could be due to a number of factors including cross core coupling, or intensity travelling through the cladding. Sect. 4.3.2.1 shows clearly that the effects are not however caused by use of the DMD, as complete modulation is achieved with a line pair width of only 3 $\times$ ICS.

4.5.2 Optical transfer function at depth

The line spread function (LSF) is useful since, from it, it is possible to obtain a 1D cross-section of the OTF. If the OTF is then known at different depths the patterns can then be chosen with a tuned spatial frequency to reject or include that depth. This section presents simulations and experimental verification of the line spread function with defocus. The LSF is then used as a mapping function between the robust and adaptable simulations, developed in Sect. 4.4.1, and the computationally efficient Gaussian convolution.

Now that the method of simulating a fibre bundle using diffraction integrals propagating from each core has been established, they are used to generate testable

results, the first of these is the edge response function (ERF), and taking its derivative, the line spread function (LSF). These are characterised at different values of defocus. The primary reason is to determine the OTF at each z distance. This is useful for a number of reasons, for widefield endomicroscopy, since we are able to determine at which depth a spatial frequency is attenuated, we are able to determine what the spatial resolution is when an object is away from the bundle, which is a useful way to report resolution and useful when near-contact imaging is desired. Furthermore, if we know the frequency transfer properties at a known depth, we can isolate high spatial frequency features for processing with approaches such as HiLo, discussed in Sect. 5.5.

Returning to SIM, from the OTF we are able to determine which modulating spatial frequencies are needed to reject features at a certain depth. i.e. if lines with a known spatial frequency are projected onto a perfectly reflecting mirror 40 μm away, and the corresponding OTF value is 50% the reconstructed intensity will be half of the in-focus intensity. That said, OTF only tells part of the story, as was seen in Fig. 4.13 bundle sampling effects play a major role, axial response profiles are needed to incorporate these. First, the LSF is simulated with defocus, then experimentally the ERF is determined by projecting the image of a knife-edge onto the proximal bundle face, then taking images at various points of defocus, and taking an intensity profile across the knife edge. The ERF is a logistic function (integral of a Gaussian), with derivative, the LSF, being a Gaussian, which can be conveniently characterised by its full width half maximum.

Simulating the ERF has a a few downsides. First, when suppressing the cladding with a Gaussian filter, the Gaussian filtering leads to an unintended smoothing of

the ERF, increase the width of the LSF. Problems are mainly manifest near focus. The mitigation approach was to choose a smoothing kernel as small as necessary for cladding suppression (75% of an inter-core spacing), and use this for the experimental smoothing kernel as well. A second problem occurred when taking an intensity profile on a line. At core boundaries on the knife edge, the ERF has a sharp intensity gradient ($\sim \infty$) and numerical curve-fitting approaches failed. The way to handle this was to take an intensity profile, aligned with the knife-edge, averaged across the columns of square of pixel values. The simulation did not reproduce the effects of global periodicity breakdown. To partially mitigate geometric bias, the bundle was randomly oriented before every OTF value was determined.

The four steps to determining the LSF at different values of defocus are shown in Fig. 4.14. The first step is to generate the knife-edge image, and the next step is to simulate projecting the image onto the proximal end of a bundle. The bundle is generated using an orientation that is random. Next, the Fraunhofer diffraction patterns are created and cast out into the Z direction, they then reflect off of a mirror, and an image is generated at the bundle face. Finally, this image is then sampled by the bundle. Diffraction patterns have different array sizes, and the process can take typically under 5 minutes for a small z distance, up to several hours for z > 250 μm diffraction patterns, ensuring that 95% of the diffraction integrals intensity was accounted for since larger fractional intensities lead to larger array operations.

The approach is, however, highly customisable, as the patterns projected from the bundle can be controlled on a core by core basis, and if desired, so can the

shape of the diffracting aperture. For the intents and purposes of SIM, however, the same pattern for each core is sufficient. As a side note, what does become apparent with defocus is edge effects. This matters when fitting the logistic curve since the fit can't include intensity decay due to the edges. The non-uniform illumination (Fig. 4.15) has brighter central cores and dimmer edge cores caused by out-of-focus light preferentially returning to the centre of the bundles as opposed to the edges.

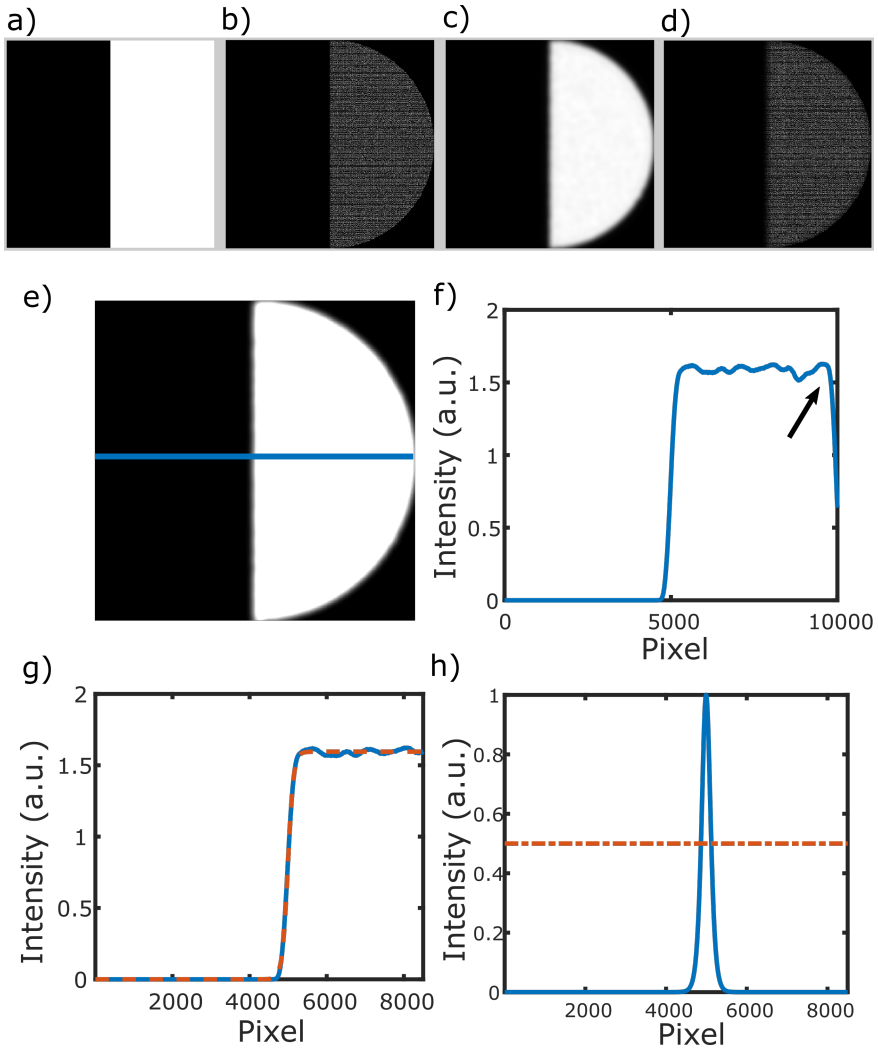


FIGURE 4.14: Edge response function simulation. (a) Edge response (ERF) image, (b) ERF sampled with the bundle, (c) diffraction integrals are propagated for each core, (d) the resulting pattern is re-sampled with the bundle, (e) cladding suppressed with Gaussian applied with kernel FWHM = 3 $\times$ ICS, (f) Edge response function from horizontal intensity profile, edge artefacts at bundle boundary identified (arrow), (g) logistic curve fit to edge response function to avoid problems with numerical differentiation, (h) line spread function with half-maximum identified.

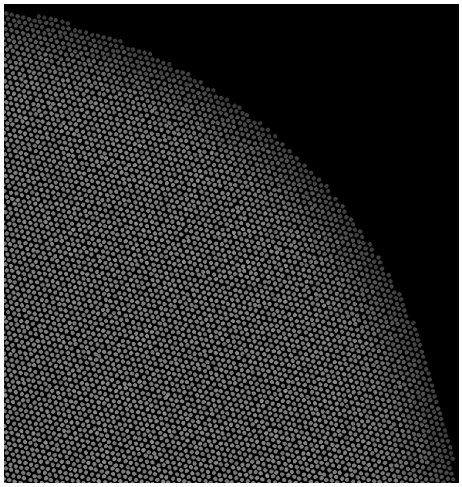


FIGURE 4.15: Edge artefacts with defocus. Z depth = 250 μm.

As a final step, the theoretical LSF from the bundle is compared to the experimentally determined one. The experiment was done by projecting a knife-edge image onto the bundle and plotting intensity profiles across the edge. The study was performed by translating a smooth metal plate, stained with a fluorescent marker, on a translation stage (Newport M-UTM25PP1HL), toward the bundle in 10 μm depth intervals and at each an intensity profile across the edge was saved, this was done until the bundle had reached contact with a smooth metal plate.

Each image was downsampled to 250 pixels, due to the inherent smoothing that was introduced, and a Gaussian filter of 2.55 μm was applied. That meant a correction factor of 960 / 250 (actual pixels size / downsampled size) needed to be applied to each measurement. The data was processed in the same way as the simulations; a logistic curve was fit to the intensity profiles, the FWHM of the

derivative then gave the LSF whose FWHM was used to parameterise it. The Z distance is reported as twice the distance reported by the translation stage due to the epi-configuration. Fig. 4.16 shows the results.

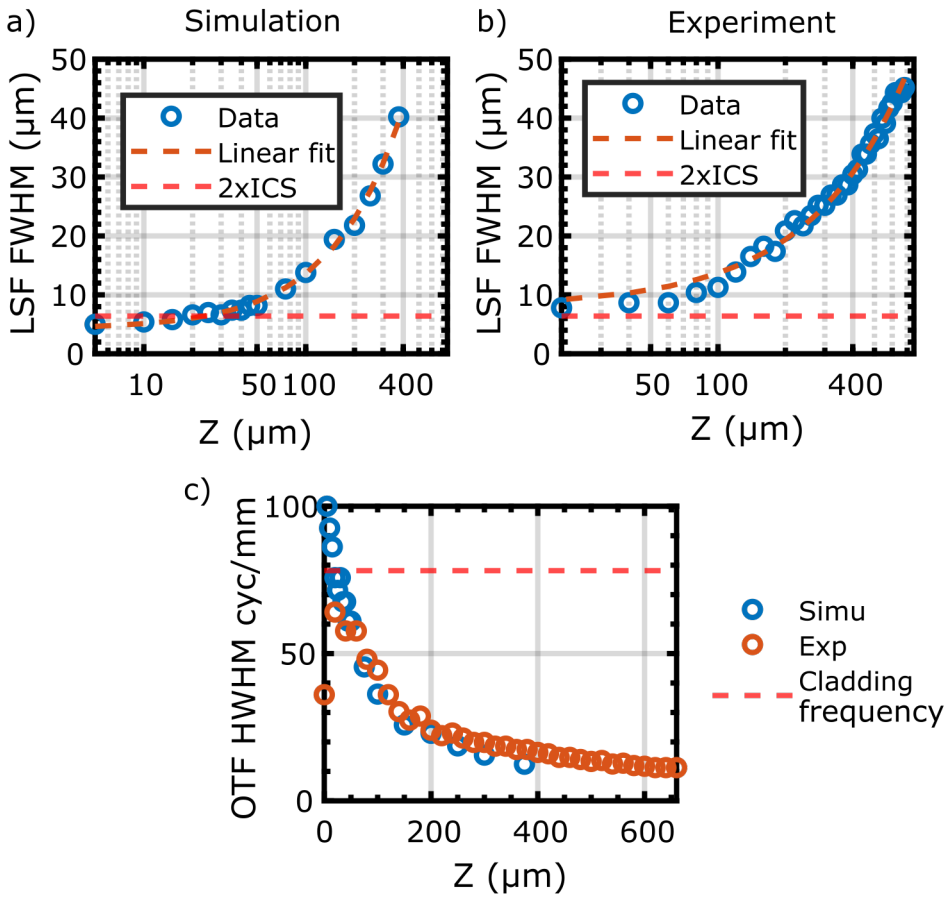


FIGURE 4.16: Line spread function FWHM as a function of distance. (a) Fraunhoffer diffraction predicted, (b) Experimentally determined, (c) Combined simulated and experimental HWHM of OTF. Intercore spacing (ICS)

The fit lines are characterised by:

$$FWHM_{Simulation} = 0.094 \times Z_{depth} + 4.2\mu m \tag{4.9}$$

$$FWHM_{Experiment} = 0.057 \times Z_{depth} + 8.0\mu m \tag{4.10}$$

As a final step, the correction factors are generated, based on experimental fitting, to the Fraunhofer approximation. The goal of this is to allow a user to generate reasonable simulations of how any arbitrary pattern may behave under defocus. The formula follows by subtracting $FHWH_{experiment}$ from $FWHM_{simulation}$, from elementary arguments what follows is:

$$FWHM_{Experiment} = FWHM_{Simulation} - (0.037z + 3.85\mu m) \tag{4.11}$$

4.5.2.1 ERF and LSF by Gaussian convolution

Simulations often don't require core-wise pattern control, and often, a more appropriate approach is to simply convolve each sampled pattern with a Gaussian to generate results for high line pair widths. By simply mapping the Gaussian convolution kernel width to the z distance corresponding the Fraunhofer simulation it's possible to use this for quick simulations. In this section, the Gaussian kernel needed to map to the Fraunhofer simulation to experimental observations at different depths is determined. An expression between the two methods is then determined via the LSF.

In order to determine the FWHM with different Gaussian kernels, a sampled knife-edge image is taken and a Gaussian filter with different kernel widths are applied. Resulting LSF's from the simulation are then mapped to LSF's for the Gaussian convolutions (Fig. 4.17). For the Gaussian Kernel MATLAB's built-in imgaussfilt.m function was used.

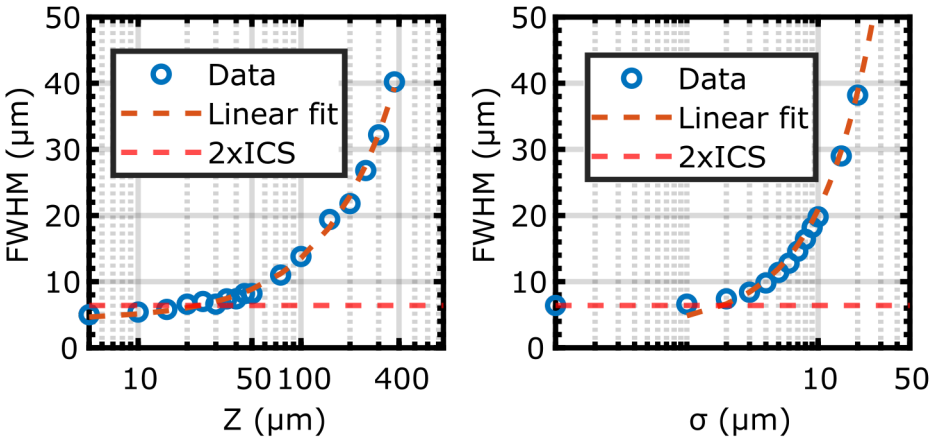


FIGURE 4.17: Comparison of LSF between (a) Fraunhoffer diffraction simulation, and (b) Gaussian Convolution. Fraunhoffer simulation was done by propagating diffraction integrals from each core of the bundle. Gaussian convolution was with filter of a certain width. 620 μm diameter bundle with 30,000 cores. Core width 3.4μm.

An analytic expression can be derived to determine the Gaussian convolution needed to simulate each depth. The expressions for the FWHM of the line spread function for Fraunhofer simulation, Gaussian convolution, and experimental are given by:

$$FWHM_{Simulation} = 0.0937z + 4.1836 \tag{4.12}$$

$$FWHM_{Gaussian} = 1.7846\sigma + 3.0592(\mu m) \tag{4.13}$$

The FWHM of the OTF tells us which modulation frequency will pass the transfer function. Since the line spacing determines the strength of the optical sectioning, the spacing must be small enough to remove out-of-focus light sufficiently [42], but large enough to pass the transfer function, and minimise the effects of pixelation and discretisation. There is another reason, introduced and characterised in the following chapter; it is for the approach to remain robust to motion.

It is also useful to know which Gaussian kernel width is useful to simulate diffraction at a known Z depth. This can be determiend by setting the equations equal to each other ($FWHM_{Simulation} = FWHM_{Gaussian}$) and solving for z and σ as follows:

$$\sigma = 0.0525 \times z_{depth} + 0.630(\mu m) \tag{4.14}$$

$$z_{depth} = 19.040 \times \sigma - 12.000(\mu m) \tag{4.15}$$

4.5.3 Axial intensity response to defocus

To determine how much out-of-focus light is rejected at each line spacing, we characterised the axial response of the system by measuring the relative intensity dropoff as a function of distance. The study was performed by translating a smooth metal plate, stained with a fluorescent marker, away from the tip of the fibre bundle at a known velocity using a translation stage (Newport M-UTM25PP1HL). The reconstruction intensity was then recorded as a function of distance by a custom-built LabVIEW VI. This study was done with a 720 μm bundle (Sumitomo IGN-08/30) as opposed to a 620 μm. The number of bundle elements is also 30,000. An example axial response profile for 32.4 μm line pair width is shown in Fig. 4.18. The curve fit shows a R squared value of 0.9997; for this reason, the axial response is parameterised in terms of the HWHM value only (Fig. 4.19). Parameterising other values, such as the 10% power-point, adds no additional information.

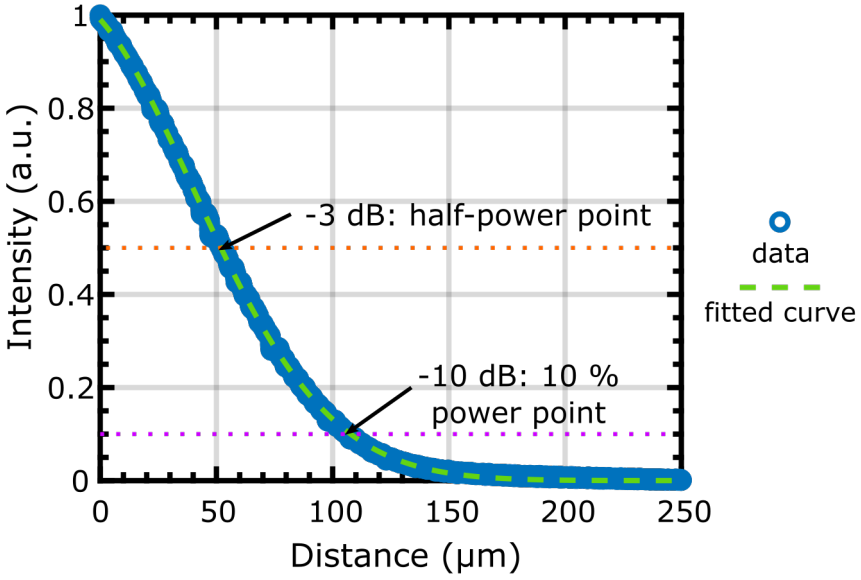


FIGURE 4.18: Axial response profile with Gaussian fit for 32.4 32.4 μm line pair width. Half-width half-max identified as -3dB dropoff and -10dB dropoff identified as well.

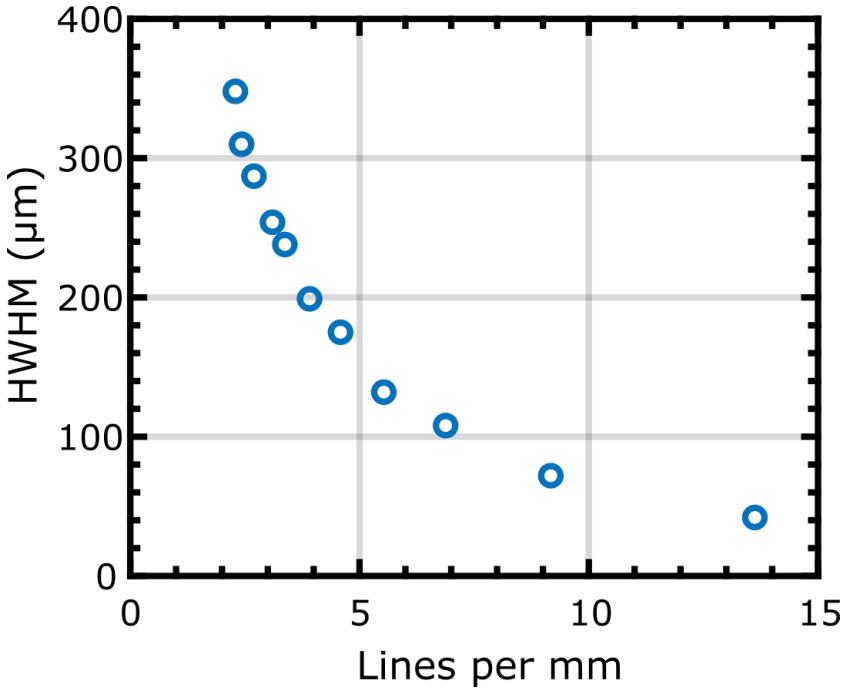


FIGURE 4.19: Parameterised axial response curve at half-power point (HWHM) as a function of line spacing. Data acquired using a smooth metal plate, fluorescently labelled with yellow highlighter, translated away from fibre bundle tip using a linear translation stage. Bundle diameter 720 μm.

4.5.4 Imaging

4.5.4.1 Tissue paper images

Tissue paper was imaged with different line-pair widths. The images were then independently normalised to 0.01% saturation; the image was iteratively scaled by a constant factor until 0.01% of the pixels were saturated. Images with line pair widths above 12x the inter-core spacing had similar intensities, while the one image taken with line pair width 17.4 μm (5.4x inter-core spacing) had a lower intensity ($\sim 30\%$ less). This agrees with the results of Fig. 4.13. The images are shown in Fig. 4.20.

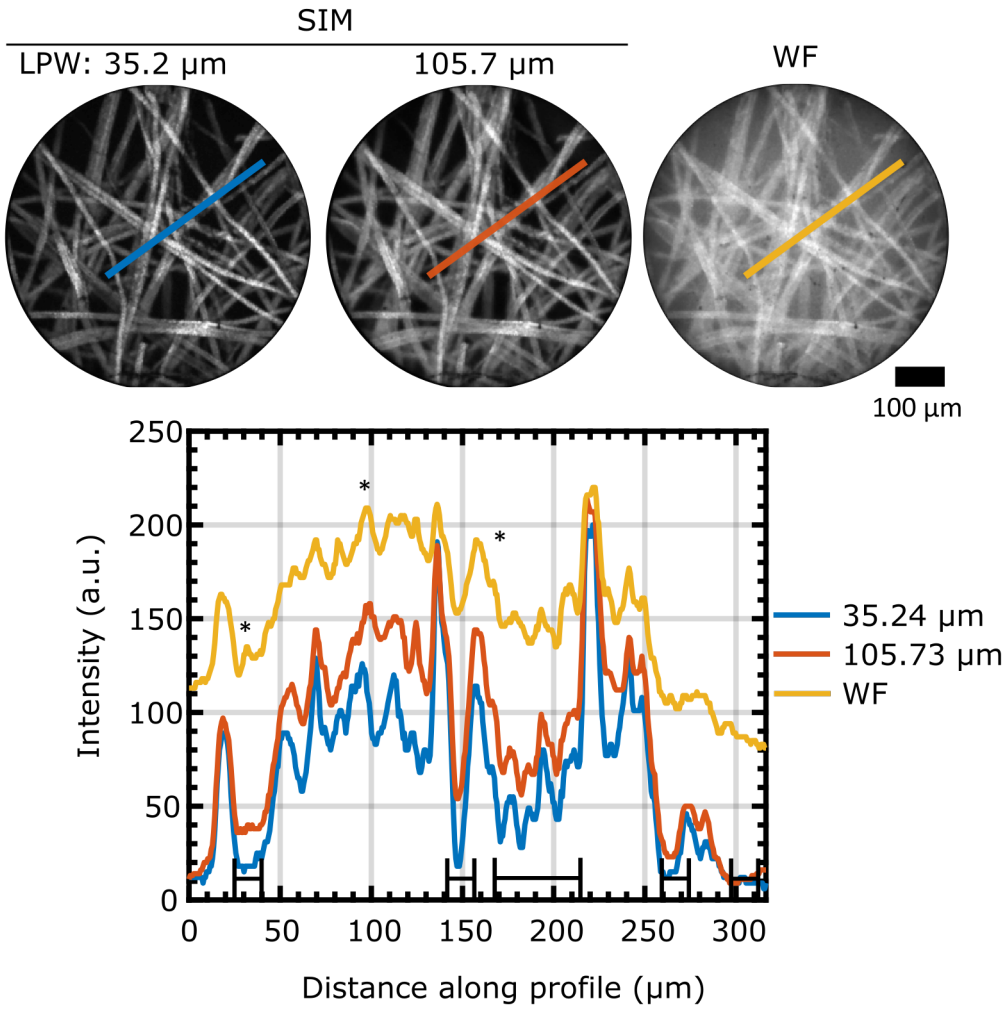


FIGURE 4.20: Lens paper fluroescently stained with yellow highlighter, intensity profile plot drawn along corresponding lines. Images independently normalised to 0.01% pixel saturation. 620 μm diameter bundle. Brackets indicate regions of clear improved far-from-focus background rejection with SIM. * indicates region where improved contrast with SIM is clear. LPW - line pair width.

4.5.4.2 Porcine stomach images

The system was validated with Porcine Stomach (Fig. 4.21). The stomach was stained with acriflavine hydrochloride (concentration 50 μg / 100 ml). The tissue was stained for 5 minutes, then rinsed with water. Images were acquired with lines of four different widths and compared to widefield images. During imaging, the probe was held steady, and only images without motion artefacts are shown. The artefacts caused by the transfer function cutting off higher frequency square wave terms, as characterised in the limitations section (Sect. 4.6) can be seen in all four SIM images. The nature of the acquisition code was to acquire three raw frames and generate SIM reconstructions in real-time. This meant that only patterns of a single modulating frequency were loaded onto the DMD at any one time. In Fig. 4.20 the probe was mounted in contact with the paper, and the patterns were changed without moving the probe. The consequence of this is reconstructions of different modulation frequencies are comparable. However, in the case of Fig. 4.21 the probe was not mounted in contact with the tissue and images of different modulation frequencies are not comparable. While each SIM image comparison relative to widefield shows improved contrast, different modulation frequencies cannot be reliably compared among tissue images.

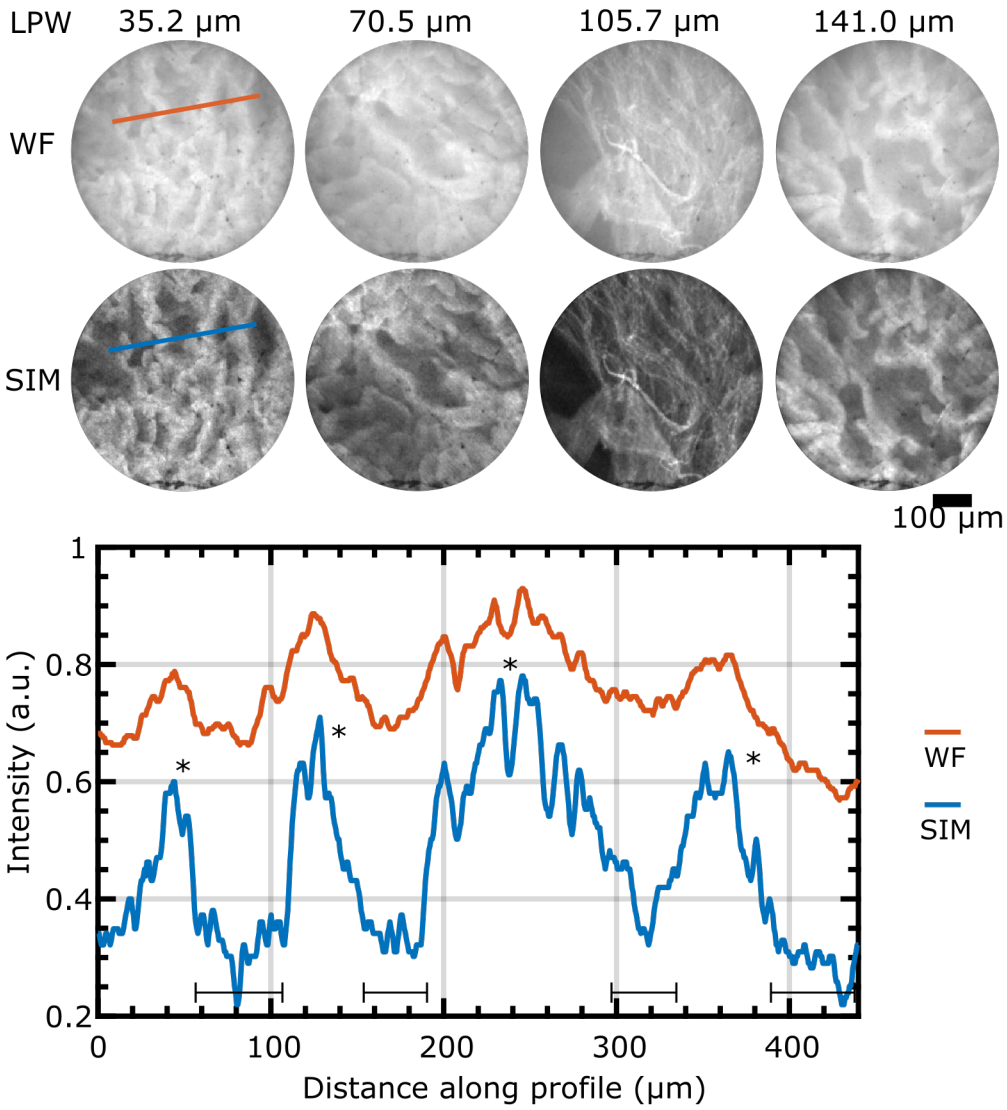


FIGURE 4.21: Porcine stomach fluorescently labelled with acriflavine. Widefield images acquired by uniformly illuminating the probe. SIM images generated at 4 different line pair widths (LPW). Profile shown for 35.2 µm LPW. (*) indicate regions of clear improved contrast, (| - |) indicate regions of clear background suppression. Images normalised to 0.01% pixel saturation independently.

4.5.5 Depth resolving

The generation of 3D resolved images at depth requires quantitative information at each depth section. The parameters required are modulation frequency, contrast transfer, and OTF, known at each depth. Conventional confocal volume

imaging involves moving a scanning spot, and lateral translation of a sample over multiple planes. The confocal spot is the same size at all depths meaning that lateral resolution is maintained with depth. In this case, the focal plane is not changed, and the 3D volume reconstructions presented here do not have a constant resolution with depth. The lateral resolution at each depth is limited by the OTF at that depth. The ability to implement axial sectioning in images, albeit with lower resolution can be adapted to the DMD configuration we present with only slight software modification. There is however a raw frame rate penalty of 3 frames per depth to be acquired. Presented now are depth-resolved images based on the subtraction-based approach introduced in anabstract by Ventalon, C. et al. (2015) [44], the method is inferred since the further manuscript is not available. Images presented were acquired with line pair frequency tuned electronically, requiring no change in a grating or physical adjustment to the optical system.

First, three images are acquired with three different line pair widths 33, 135, 400 μm. These have cutoff depths of 60, 400 and 1000 μm, and modulation depths 0.45, 0.9, and 0.9. We then image tissue paper($\sim$ 50 μm) separated from a fluorescently labelled piece of paper by a $\sim$440 μm plastic sheet. Three images are constructed, the first is just the SIM reconstruction at 33 μm, the second, the difference in SIM images acquired at 135 and 33 μm, and the third, the difference in SIM images acquired at 135 and 400 μm. In-depth the images correspond to $\sim$ 0-60 (the axial slice of the lens paper), 60-400 (in the plastic bag), and 400-1000 μm (the fluorescent background). The three regions are very clearly isolated in Fig. 4.22.

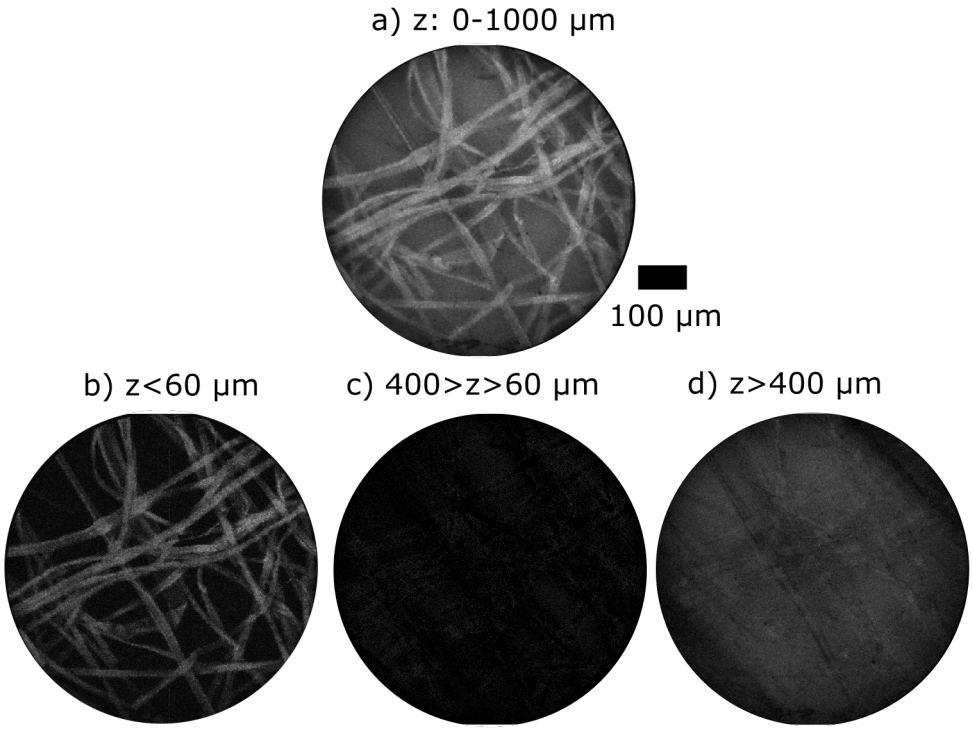


FIGURE 4.22: Fluorescently stained lens paper ($\sim$ 50 µm) separated from a fluorescently labeled piece of paper by a 440µm plastic bag. (a) axial slice 0 - 1000 µm, (b) 0-60 µm, (c) 60-400 µm, (d) 400-1000 µm

4.6 Limitations

4.6.1 Spatial phase misalignment

The spatial phase relationship of the sinusoidal or square wave patterns needs to remain aligned or artefacts are introduced. The spatial phase misalignment artefacts have different characteristics depending on the type of modulation (sine waves vs square waves). A constant spatial phase difference between raw frames could occur, for example, when using a motorised grating with poorly syncronised external triggering. If an external camera trigger pulse introduced an added

delay between each frame this would result in patterns which had advanced further than their intended position, and hence, artefacts. These aretefacts are identical in nature to those that occur with a probe moving at a slight angle to line orientation. This is briefly shown in (Fig. 4.24) and quantified in Sect. 5.3.2. To study the effect of a spatial phase shift between frames sine waves are used with the same spatial frequency as the spacing in a square wave pattern. The study is done with a high precision mesh, and the mean of the absolute deviations is reported for both sine and square waves.

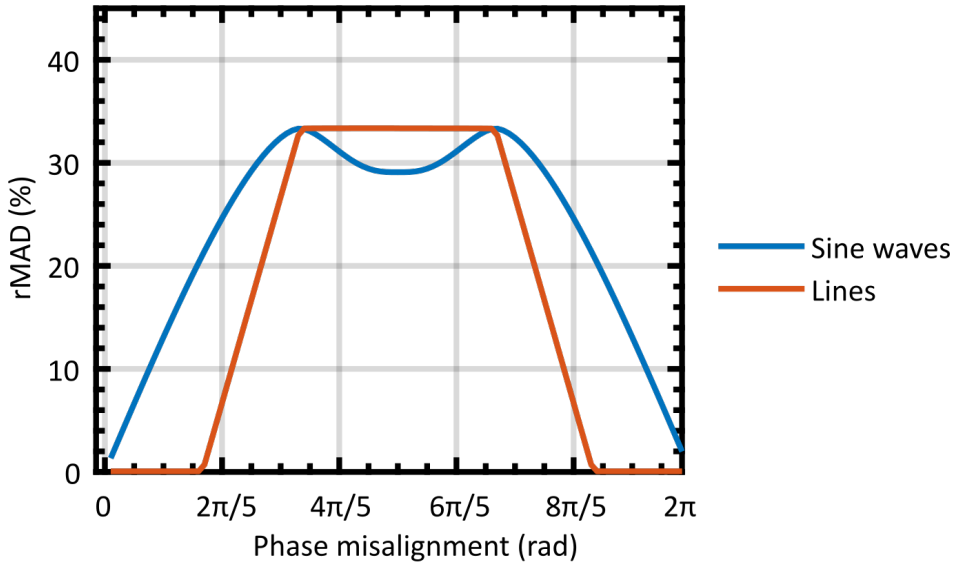


FIGURE 4.23: Percent error with constant phase misalignment between SIM frames.

Several things are evident from Fig. 4.23. Firstly, that square waves are more robust to small misalignments. The reason for this has to do with the nature of the reconstructions. Consider a single-pixel anywhere along the line; perfect reconstructions are obtained when in a sequence of three images the pixel has intensities [0,1,1] or [0,0,1], a small shift only leads to the replacement of [0,1,1] with [0,0,1] or vice versa. Meaning the reconstructions are still perfect, even with

misalignment. The phase misalignment at which in-focus reconstructions are no longer perfect is $\pi/3$. This robustness of line patterns to spatial phase misalignment further seems to indicate that when an image is in focus, this will not contribute residual line artefacts.

Phase misalignment has been described before Ref. [138, 139], sometimes referred to as phase mis-stepping. The common reason this occurred was due to incorrect shifts in a motorised grating or piezo device. A DMD does not suffer from this phase misalignment in the same way, however, that's not to say the configuration is immune. The exact same artefacts occur with a moving probe. The method of motion compensation (Chap. 5) will discuss the exact mechanism, but in brief, Fig. 4.24 shows a circular viewing window capturing a stationary feature with a circular window that shifts about 1/4 of its diameter between acquisitions. If the circular viewing window moves perfectly parallel to the orientation of the lines, and the exact motion is known, reconstructs are perfect, if the motion and orientation vectors are offset a spatial phase misalignment is introduced between the lines. The phase offset can be clearly seen in the unequal white and dark areas in the overlapping aligned phase, and misaligned phase images.

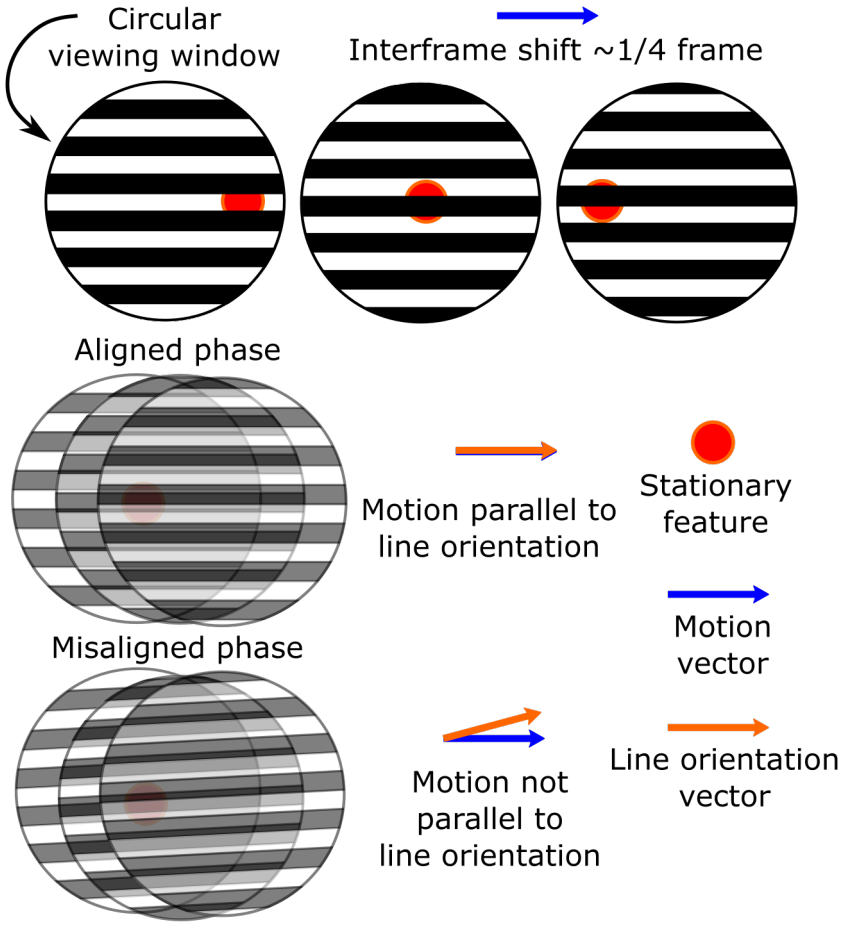


FIGURE 4.24: Effect of motion on phase misalignment. A circular viewing window, with well known motion, shifts between three acquisitions of a SIM sequence. When motion vectors and line orientation vectors are parallel there is no phase misalignment.

4.6.2 Residual line artefacts

Residual artefacts result from using square wave modulation instead of sine waves. These have previously been reported as Trapezoidal artefacts [138, 139]. When evaluating the performance of lines as opposed to sine waves there are three cases that need to be considered: (1) the in-focus plane, (2) the far from focus plane, (3) intermediate planes. The in-focus case has been shown, the far from focus case is obvious since substituting a zero into Eqn. 4.2, simply yields 0 intensity. The third case, however, is far more subtle. Fig. 4.25 shows the reconstructions of sine

waves and line pairs when the plane is only slightly out-of-focus. This means that more than 2 Fourier terms pass the pass the OTF but few enough that the modulation becomes trapezoidal.

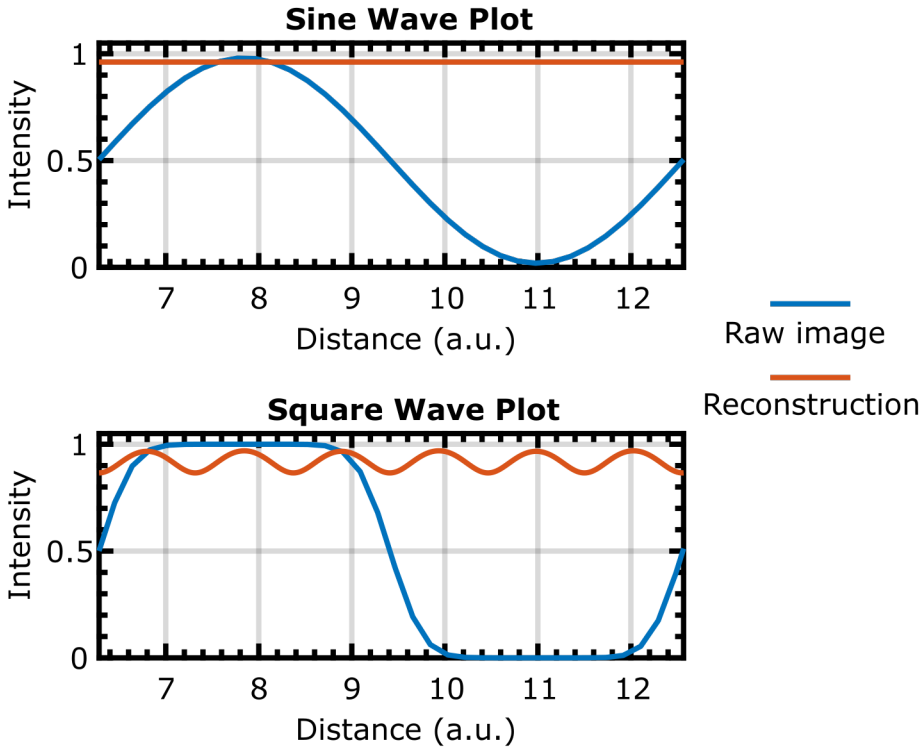


FIGURE 4.25: Defocus artefacts for sinusoidal and line modulation at axial response drop-off of $\sim 10\%$.

By decomposition of line patterns into a Fourier series and rejection of higher-order terms, the residual artefacts can be seen to be a property of defocus, not a property of misalignment.

Square waves can be decomposed into elementary sinusoidal functions using the Fourier series:

$$I_{cs} = (\frac{4}{\pi}) \times \sum_{n=1}^{N} \frac{1}{n} \sin(\frac{2\pi n x}{P} - \phi) \tag{4.16}$$

where I_{cs} is a one-dimensional cross-section perpendicular to the modulation, N is the number of terms in the sum, P is the spatial period, ϕ is the phase. If

we make the simplifying assumption that the OTF acts as a unit step function with a clear cutoff, that's equivalent to generating lines with fewer terms. If we know the OTF cutoff at each position of defocus, as is simulated and determined experimentally in Sect. 4.5.2, then we know the number of terms that are rejected by the OTF. A simulation was run with varying numbers of Fourier terms to quantify the effect of residual artefacts. The mean of the absolute differences was used to characterise the quality of the reconstruction. Fig. 4.26 shows an example reconstruction. The intensity falls below zero at a number of points, and may on first glance appear non-physical. There are a number of simplifying assumptions which cause this behaviour. Firstly, a typical system would have modulation depths less than 1, second, additional terms in the Fourier series compensate for this behaviour, these are often damped, but in these reconstructions, we have truncated them. Similarly, the 3rd term in the series is included undamped.

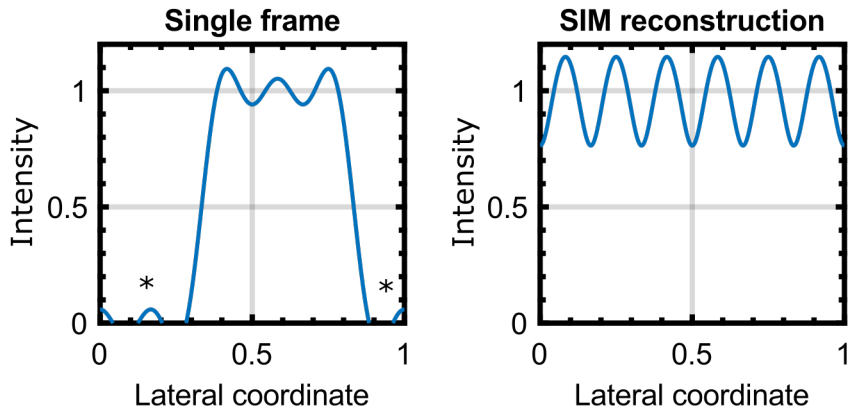


FIGURE 4.26: Simulation of square wave reconstruction using N=3 Fourier terms. (*) indicate regions > 0.

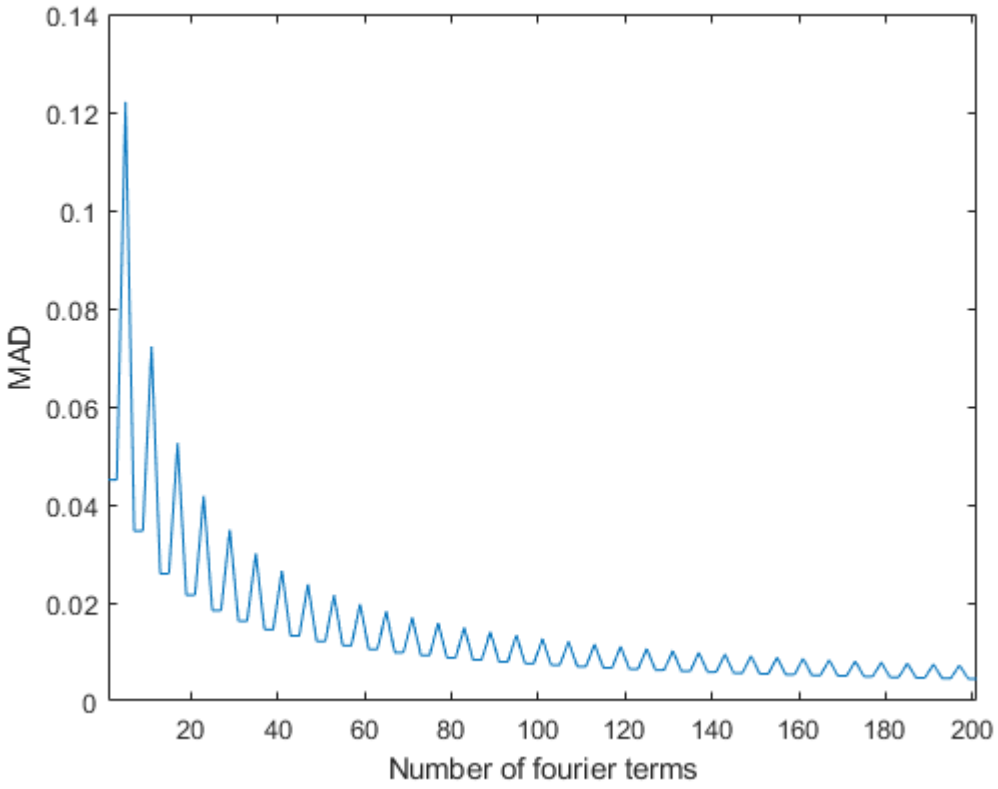


FIGURE 4.27: Relative mean of the absolute differences, rMAD, as a function of number of Fourier terms (0.10 = 10% difference). Correction factor used was for square waves which explains the non-zero behaviour at N=1. Total number of terms including even (200 total terms = 99 odd terms).

Quantitatively, the residual artefacts as a function of line pair width are quantified using the relative mean of absolute differences (rMAD) in Fig. 4.27. One valid question is that when N=1, only the first term in the sine wave is included, and we would expect the MAD to be 0. So why isn't it? The reconstructed images are compared to an ideal SIM reconstruction using lines which means the scaling factor is $\frac{1}{\sqrt{2}}$, yet the sine scaling factor is $2\frac{\sqrt{2}}{3}$. In reality, the reconstructions are complete, albeit no longer quantitatively correct. It may look on first glance that the values could simply be divided by the lower envelope to obtain a metric which only characterised residual artefacts. The perfect reconstructions, however, only holds for N=1,3, and tail values. Residual line artefacts are visually

observed in **all** other N values. Perhaps most interestingly, these artefacts have other properties as well, such as the fact that their spatial frequencies are mixed and not the same as the modulating frequency.

The only step required to simulate defocus artefacts with any arbitrary system is to use the measured OTF at different values of defocus. The number of Fourier terms can be generated until the OTF cutoff is reached.

Lines projected onto a fluorescent metal plate experimentally show the properties of residual artefacts. By choosing a modulating frequency much lower than the transfer function, and taking its FFT, it is shown that the fundamental Fourier frequency and the higher-order terms (14 total terms) are within the transfer function of the system. The metal plate is then moved away from focus to show defocused lines and their FFT. The FFT shows the impact defocus has on the transfer function and shows the rejection of higher-order terms. Next, the reconstructed images in-focus and out-of-focus are shown, and the artefacts are characterised using the FFT. The modulation frequency is demonstrated not to be the same as the artefacts. For the avoidance of all doubt, the artefacts are also shown to have a different spatial frequency from the pixel pitch. The results are shown in Fig. 4.28.

What is surprising is that defocus artefacts occurred in both the in-focus and out-of-focus case. In both cases, the spatial frequency of the artefacts is not the same as the modulating spatial frequency. Analysis of the two frequencies showed that the fundamental frequency of the artefacts was the same; it is a factor of 3 higher than the modulation frequency. This can be demonstrated by truncating Fourier sine terms. The Fourier transform reconstructions of square waves were used to build up to two SIM reconstructions to exact match to properties

of the transfer functions. For the in-focus reconstructions, 14 odd terms are used (28 total), and for the out-of-focus reconstructions I use 5 odd terms (10 total), the reconstructions are shown in Fig. 4.29. They show the same behaviour we are experimentally observing in Fig. 4.28. This confirms that defocus artefacts, are not only a bi-product of misalignment but are in fact a product of higher-order reconstruction terms failing to pass the OTF leading to incorrect reconstructions.

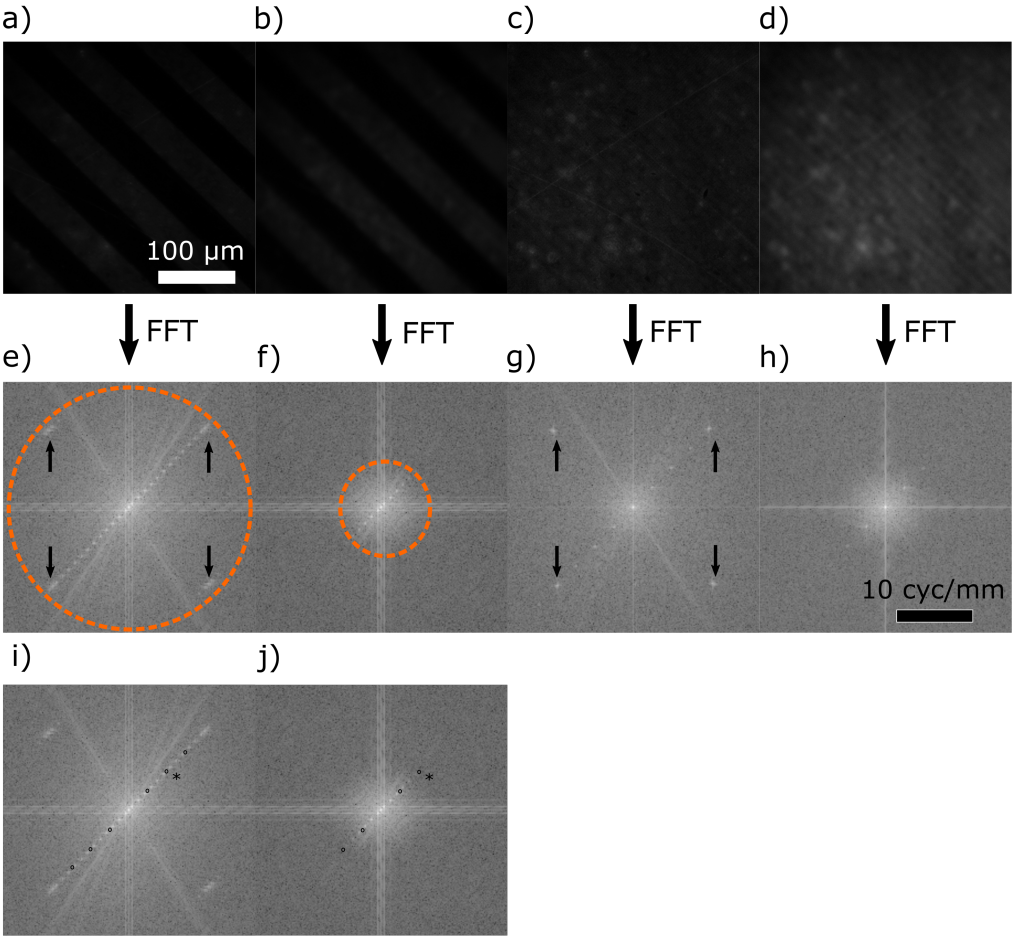


FIGURE 4.28: Reconstruction showing residual line artefacts of SIM reconstructions. Raw pattern with (a) in focus lines, (b) defocused lines, (c) in-focus SIM reconstruction, (d) defocused SIM reconstruction, (e-h) corresponding FFT's for (a-d), (i,j) FFT of lines with locations of reconstruction artefacts showing that reconstruction artefacts are not at the same peaks as the modulation. Circle - OTF, arrows indicate terms resulting from pixel structure.

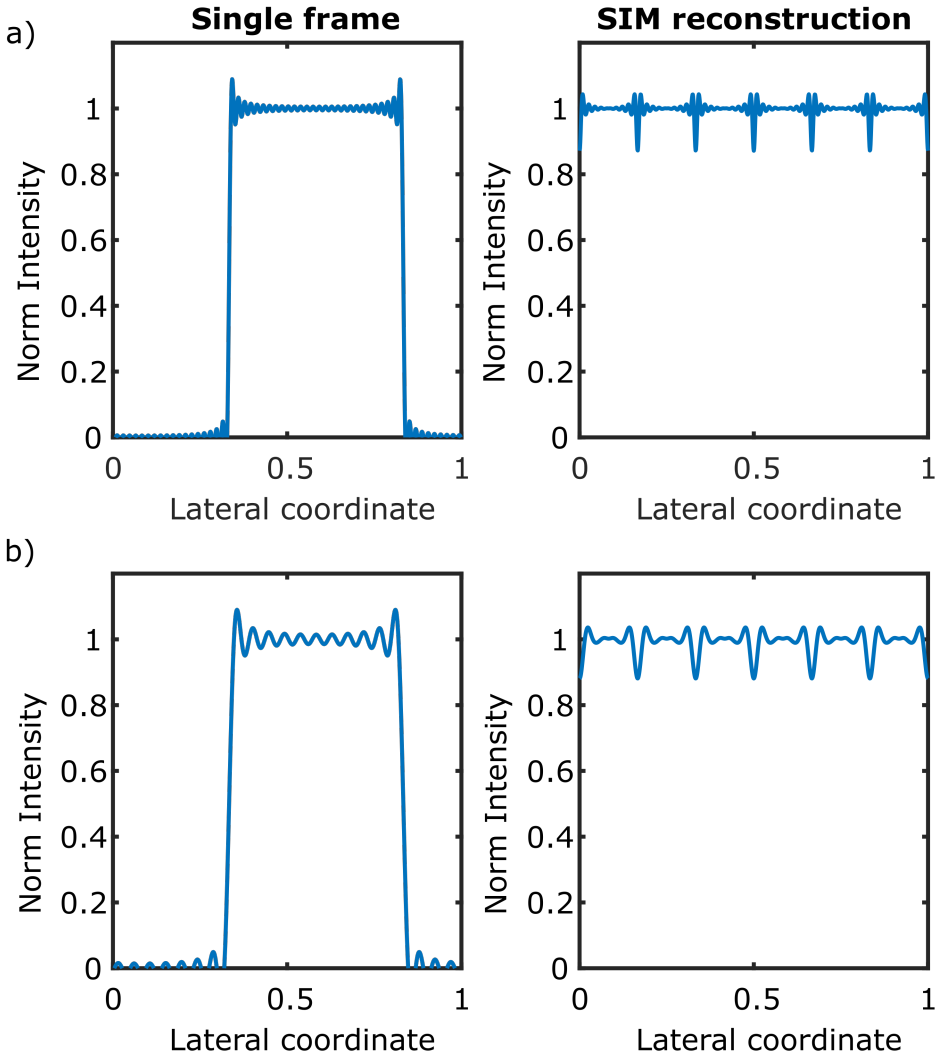


FIGURE 4.29: Comparison of Fourier reconstructions of square waves using 14 odd terms, meant to correspond to in-focus reconstruction in Fig. 4.28(e), and 5 odd terms, meant to correspond to out-of-focus reconstruction in 4.28(f).

In conclusion, the penalty we pay for using square wave modulation rather than sinusoidal is that demodulation does not perfectly hold at intermediate stages of defocus. This penalty is only paid on light returning from near focus, at points where the axial intensity response has not attenuated by more than $\tilde{2}5\%$. The simulation is not perfect, the assumption that the OTF acts as a cliff's edge, and higher spatial frequency terms are omitted when in reality they are damped and this may have an effect on the sharp peaks in Fig. 4.27, it may even lead to intermediate values, bounded by the envelope. What's more, as discussed, the

correction factor at N=1, should be the sine scaling factor, which would make the MAD = 0 for N=1,3 (areas with perfect reconstruction). The scaling factor for intermediate values between N=1 and N=∞ isn't easy to determine due to lack of perfect reconstruction at any intermediate values, and the inability to fit a horizontal line due to lack of symmetry.

It may appeal to an operator to have a grid pattern perfectly conjugate to a focal plane for imaging. What the MAD of reconstructions at defocus suggests is that patterns should be blurred so that at the object plane, there is only the fundamental frequency. This is only the case so long as the contrast is maintained. Line pair defocus is not the only source of artefacts, and more are discussed in the next section.

4.6.3 Motion artefacts

At this point, it may be tempting to attempt to generate mosaics by stitching adjacent frames together. Especially because endomicroscopes are used in several configurations, including through the working channel of an endoscope[145], as part of a robotic or mechanical system[45, 146], or in freehand operation[30]. However SIM is a three frame process, and motion between frames leads to artefacts, these artefacts are different depending on whether or not the probe is aligned with the lines or perpendicular to them (Fig. 4.30). It is for this reason that mosaics have not, previous to work in the motion compensation chapter, ever been reported. In the following chapter, a new approach to mosaicking is presented.

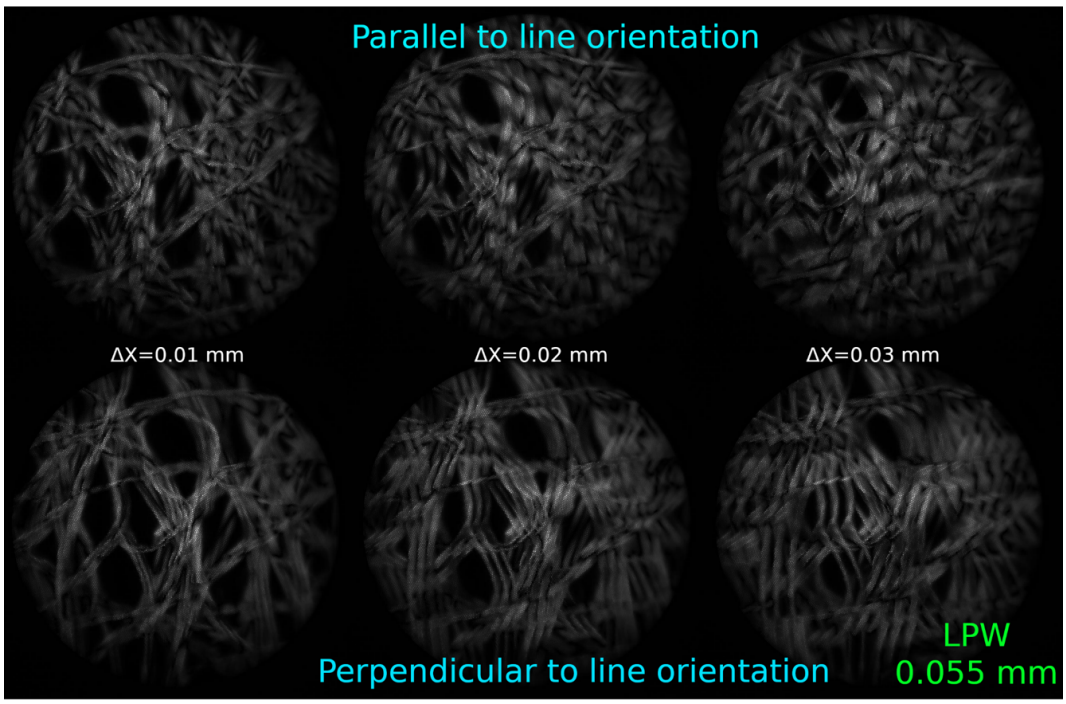


FIGURE 4.30: Motion artefacts in a single image when a probe is moving relative to the lines. 720 μm probe. Shift between frames stated as Δx.

4.7 Conclusion

This section presented optical sectioning in endomicroscopy via the first known approach which uses SIM and a DMD. The work in this chapter was done with lines that were aligned with the rotated rectangular array structure of the DMD, reducing discretisation artefacts. A parameter set for optimal contrast sectioning, minimum artefacts, as well as optimal modulation depth was developed theoretically and experimentally. The OTF with defocus was reported, as was the axial response to defocus. The cause of line artefacts was simulated and shown experimentally. Improved imaging was shown for lens paper, and validated with *ex-vivo* stained bovine stomach. The depth sectioning capability was shown for lens paper separated from a fluorescent background.

Chapter 5

Structured illumination with motion compensation

5.1 Introduction

Endomicroscopy is often used in the working channel of an endoscope, or in free-hand operation. There have also been preliminary investigations into using endoscomicrocopes mounted in mechancal or robotic scanning systems [146–148]. These three configurations all tend to involve a probe which is constantly, or often, in motion. In some circumstances, motion is deliberately introduced. Mosaics are often useful since a conventional biopsy is typically around 3 mm x 3 mm in size, while a fibre endomicroscope has a typical probe diameter of less than 1 mm (Fig. 5.1). This size difference can make it difficult for clinicians to get a broader view of the tissue morphology[149]. However, stitching together neighbouring frames through a mosaic can increase the effective field of view (FOV) [14, 78, 120, 121, 149], mosaicking is discussed in the technical methods (Chap. 3). While mosaicking is useful, SIM is a three frame acquisition process and motion between acquisition frames makes the reconstructions susceptible to artefacts. The artefacts appear as blurring or chopping, as was shown in the SIM chapter, explaining why, before the work in this chapter which led to an oral presentation

at Photonics West and a publication [17, 46], mosaicking in structured illumination endomicroscopy had not been previously demonstrated.

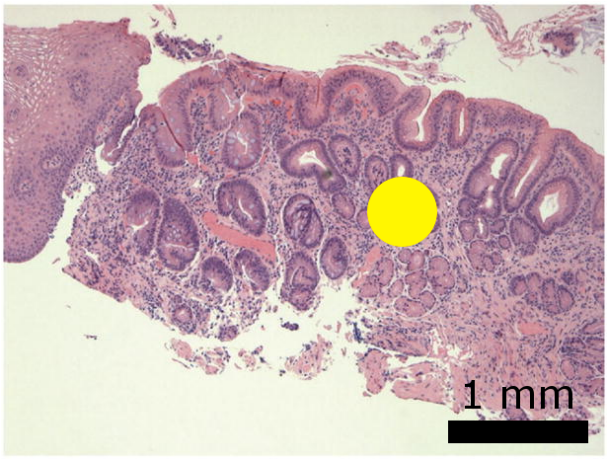


FIGURE 5.1: Image showing the relative size of a typical histology and a fibre bundle (yellow circle). Mosaicking is useful to increase the effective area of the probe.

This chapter reports motion artefact compensation and mosaicking of SIM endomicroscopy images. As will be discussed in more detail in the following sections, there are two steps involved in this: registration correction, and alignment correction. The first step is to register the images. However, the three images used for SIM have a superimposed modulation pattern and it is not straightforward to register the images directly. To address this, two wide-field frames are introduced which book-end the three-frame SIM sequence. A 5 frame sequence is then captured, register the shift of the two wide-field images and estimate the shift between SIM images.

However, image registration alone isn't sufficient to correct for motion artefacts since the line-pairs in the image will not be aligned. To align them the second step is to take the direction of the probe, which is determined as part of the registration correction, and reorient the line patterns to the direction of travel.

The remainder of this chapter describes and evaluates this new approach to motion compensation for SIM endomicroscopy, demonstrating real time mosaicking when using a translation stage or freehand probe.

5.1.1 System description

The same system, the one previously reported in the SIM section (Sect. 4), is used for both mechanical and automated motion compensation with the exception of HiLo in Sect. 5.5. Fig. 5.2 shows a 3D representation. The DMD model is not identical to the one used (DLP 2000 shown, instead of 3000), however, it has approximately similar dimensions. Light from the blue LED of the DMD's built-in illumination system is directed through a 20x/0.25 NA finite-conjugate objective (Newport), a 20 mm collimating lens (Thorlabs ACL2520U), an excitation filter (Thorlabs FESH0450), and reflected off of a 490 nm cut-on dichroic mirror (Thorlabs DMLP490) onto the back aperture of a 4x / 0.25 NA finite-conjugate objective. It is then transferred through a fibre optic bundle (Sumitomo IGN-08/30) in pixelated form. The returning fluorescence emission is imaged onto a camera (Point Grey Flea3) via the 4x objective, dichroic beam splitter, and an emission filter (Thorlabs FELH0500). A microcontroller (Pyboard 1.0, not shown) is used to control the triggering of the camera and DMD. A data acquisition (DAQ) card (NI USB 6008) provides communication between the controlling PC and the micro-controller. At the object plane, each camera pixel corresponds to 0.336 μm x 0.336 μm, with the 4 μm intercore spacing of the bundle this provides better than Nyquist sampling. Each DMD pixel corresponds to 4.407 μm x 4.407 μm measured at the object plane. The power at the fibre tip was measured to be 95 μW,

and the fibre coupling losses were measured to be 89%. The system has no optics on the distal end of the fibre, meaning images are acquired in contact mode only.

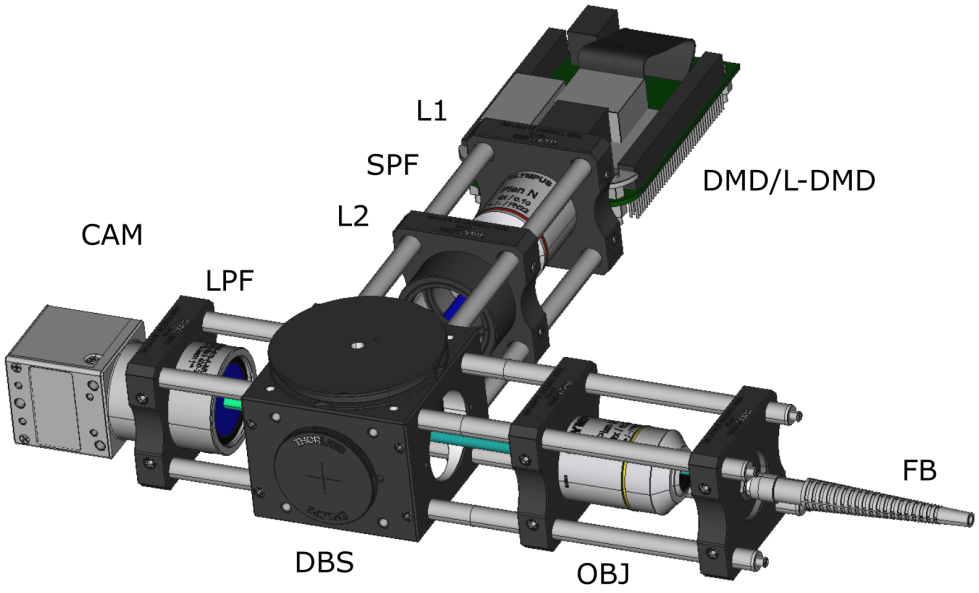


FIGURE 5.2: Schematic of endomicroscope. DMD - Digital micro-mirror device, OBJ 1 - 20x, OBJ 2 - 4x objective, L1 - 20 mm collimating lens, L-DMD - lens built into DMD, SPF - Short pass Filter, LPF - Long pass filter, DBS - Dichroic beam splitter, FB - fibre bundle.

5.2 Compensation for mechanical motion

In some emerging endomicroscope configurations, such as when the probe is used in a mechanical arm, the direction and magnitude of the inter-frame shift are known *a priori*. SIM, in the way employed here, involves a one-dimensional modulation pattern, so if the probe shifts transverse to the modulation, as long as the images are registered and shifted prior to applying the SIM images, it is perfectly possible to reconstruct SIM images. Mechanical motion is a particular application of more general motion compensation, for which the detailed steps of registration and orientation correction are covered later in Sect. 5.3.3.

In order to test the feasibility of motion compensation, a simulation of the mechanical translation of a probe is designed. This required a calibration step, and a registration step. The calibration only needs to be done once.

5.2.1 Calibration

The direction of the probe needs to be determined using a translation stage moving at a constant velocity. To do this, a number of widefield images at various translation stage positions are captured and for each step the direction of motion was recorded. Once the direction of motion with respect to the image on the camera is known, lines were generated on the DMD which were parallel to the motion, meaning the modulation was transverse to the motion. The probe is operated in contact mode and when imaging tissue paper on a translation stage, the probe remained in static contact for ~50 μm before gliding over the tissue paper.

The interframe shift (IFS) and direction of motion are shown in Fig. 5.3. The mean IFS was 0.59 μm, the relative MAD was 1.25 % (7.46 nm), the mean direction was 0.79 rad, and error was 0.11 % (0.00066 rad, 0.037 deg). These error values are practically negligible especially when considering the inter-frame shifts on the order of 50 μm, and alignment sensitivity is on the order of 5-10 degrees. Error is also non propagating as the translation stage works on changing the absolute position of the stage, not the relative. The orientation of the patterns was adjusted to be parallel to the direction of motion.

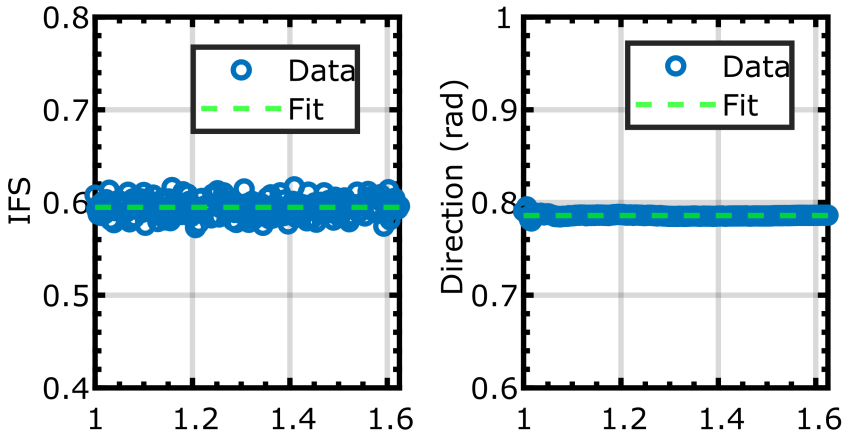


FIGURE 5.3: Inter-frame shift and direction. Probe mounted in contact with fluorescently stained lens paper mounted on translation stage.

5.2.2 Registration

Now that the motion direction has been calibrated, the images are registered. To eliminate any error which may come from small deviations in the stage position, the position reported by the stage is not used. Instead, the book-end wide-field images are used to register and estimate motion. Doing this allows us to determine the magnitude of motion, as well as the direction, which simply matches the calibration direction. The revised frame sequence is shown in Fig. 5.4. the magnitude of the shift of the individual images is estimated as 1/4 the shift of the widefield images.

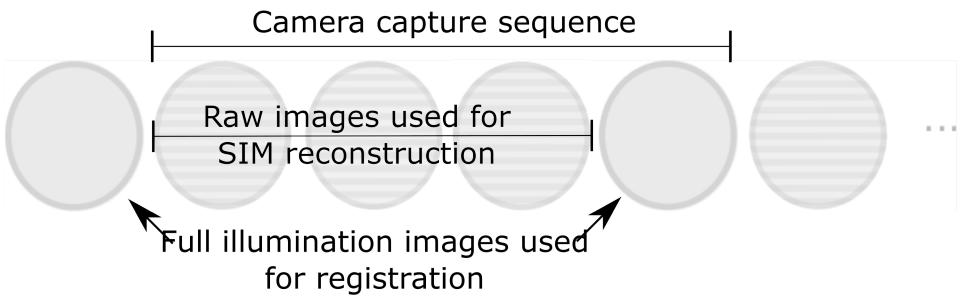


FIGURE 5.4: The four-frame acquisition sequence. The three SIM frames, with shifted line patterns, are book-ended by widefield illuminated frames used for image registration.

5.2.3 Validation and preliminary results

To quantify the improvement in a practically meaningful sense, using lens paper, the stage is translated laterally across the distal face of the probe. The preliminary results, presented at Photonics West in 2018, are shown in Fig. 5.6. Besides clear qualitative improvement, the improvement is quantified by comparing the reconstructions to statically obtained reference images using the maximum value of the normalized cross-correlation matrix (NCC). The NCC method is discussed in Chap. 3.

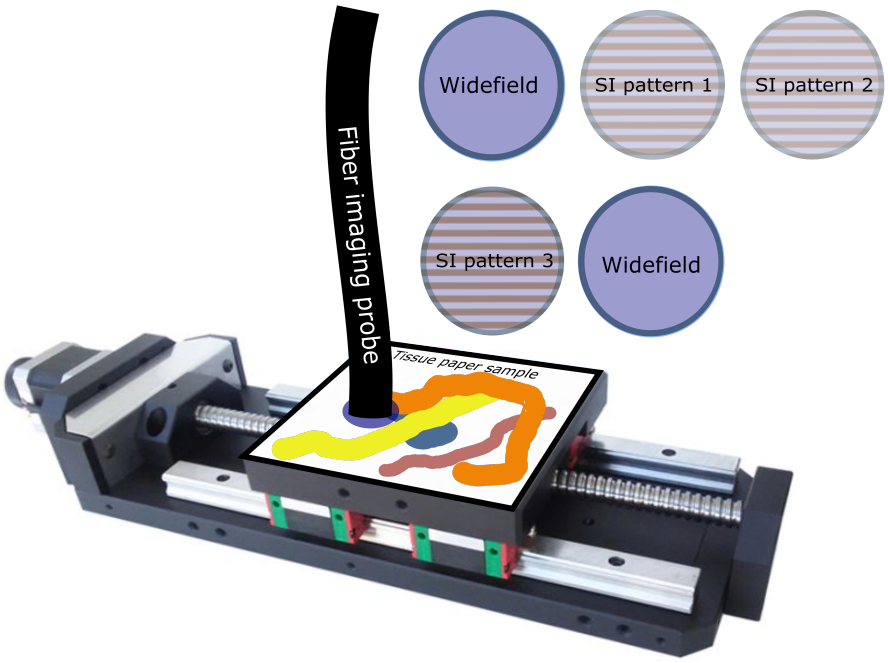


FIGURE 5.5: Motion compensation approach with a translation stage.

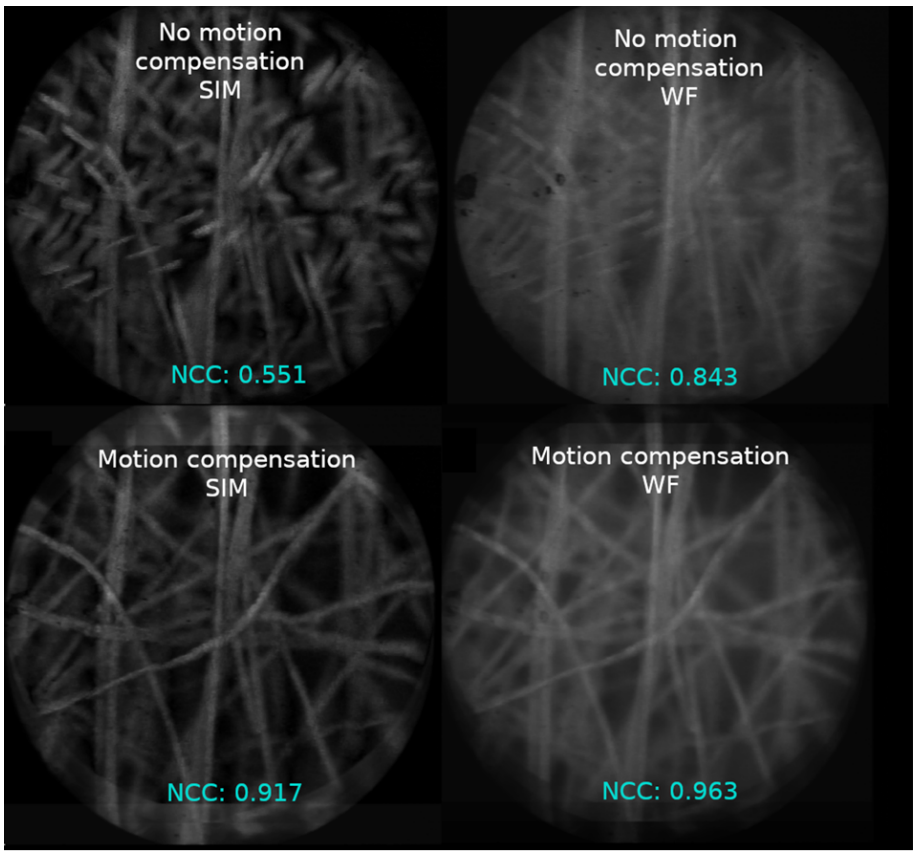


FIGURE 5.6: Images with and without motion compensation. NCC - normalised cross correlation value when compared to a stationary reference. WF - widefield reconstruction ($I_1 + I_2 + I_3$).

A few key conclusions can be drawn from these results. Qualitatively, there is a clear improvement in both SIM and widefield reconstructions including complete reduction of blurring and chopping. Quantitatively, the NCC value, when images are compared to a stationary reference improves from 0.55 to 0.92 with SIM on, and from 0.843 to 0.963 with widefield reconstructions. This metric is commonly used in mosaicking indicating this has the potential to lead to mosaicking.

The applicability of this approach is limited, since the direction of motion of a probe often isn't known in advance. It is more useful to have an approach that can automatically compensate for motion in real time.

5.3 Compensation for freehand motion

A scheme to automatically correct for motion, including freehand, in real time is now introduced. When compensating for mechanical motion, during the registration step both the magnitude and direction of motion were determined. Previously, only the magnitude of motion was used to reconstruct images. Using both magnitude and direction information to shift the images relative to each other SIM images are reconstructed. This introduces an additional problem, that the lines are once again misaligned. To correct this, a feedback-based approach is introduced, where, at the end of each frame sequence once the direction is determined, the orientation of the lines is updated parallel to that direction, in real-time.

Since freehand motion is different from mechanical, there are a few assumptions that need to be validated. First, the motion between the frames must be approximately linear. Second, the shift between the frames must be approximately uniformly spaced.

5.3.1 Probe trajectory and motion characterisation

A trajectory of the movement of the probe is generated to verify the frames are uniformly spaced and approximately linear. This is done by capturing widefield images with a probe being translated across a fluorescently stained lens paper sample, registering the images, and plotting the path of the probe.

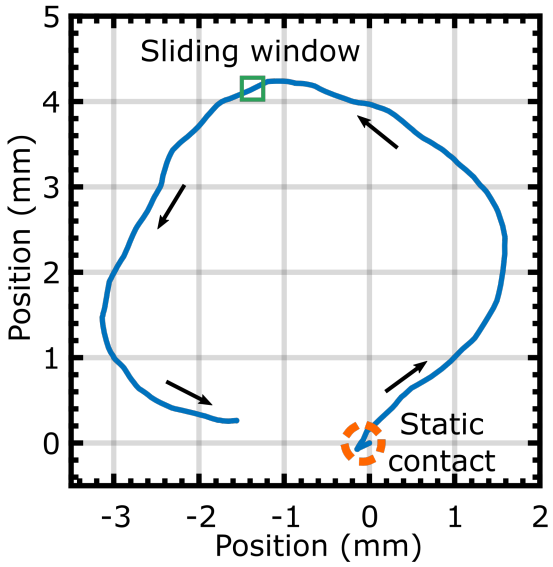


FIGURE 5.7: Trajectory of a probe being translated circularly over lens paper. Arrows indicate the direction of the probe. Sequential images are registered using the maximum of NCC matrix. The sliding window indicates a region enclosing only 5 points which moves along the direction of the trajectory. Static contact indicates the probe trajectory before the probe is in constant motion.

The probe trajectory is shown in Fig. 5.7. Neighbouring sets of 5 points are isolated by sliding a window along the trajectory of the probe (Fig. 5.8) and interpolate the ends. To verify uniformity, the average distance between frames is reported and compared to the mean. To verify linearity, the Euclidean distance is calculated, the distance between the interpolated line and the probe position (EucD). The error from the expected position is computed, by determining the location of three evenly spaced points on the interpolated line and calculate the distance from the actual probe positions. Frames are captured at the full frame rate of the camera (120 frames/s). However, to compare this to real time system performance (10 frames/s) the data is downsampled to 10 frames/s, and report this alongside high speed acquisitions in Tab. 5.1.

	IFS (μm)	μD (μm)	μD (%)	EucD (μm)	EucD (%)	ExpD (μm)	ExpD (%)
TS	19.3	0.9	4.4	0.6	2.9	1.2	6.1
TS_{ds}	77.0	1.6	2.1	1.4	1.8	2.5	3.2
FH	8.6	1.1	16.8	0.3	4.0	1.4	21.2
FH_{ds}	77.5	19.3	27.2	2.4	3.3	22.7	32.9

TABLE 5.1: μD - mean deviation from uniform motion, EucD – mean Euclidean distance from interpolated line, ExpD - probe distance from expected position. TS - translation stage, FH - freehand, ds – downsampled. Neighboring sets of 5 coordinates. Values in bold are less than the resolution of the probe (6 μm).

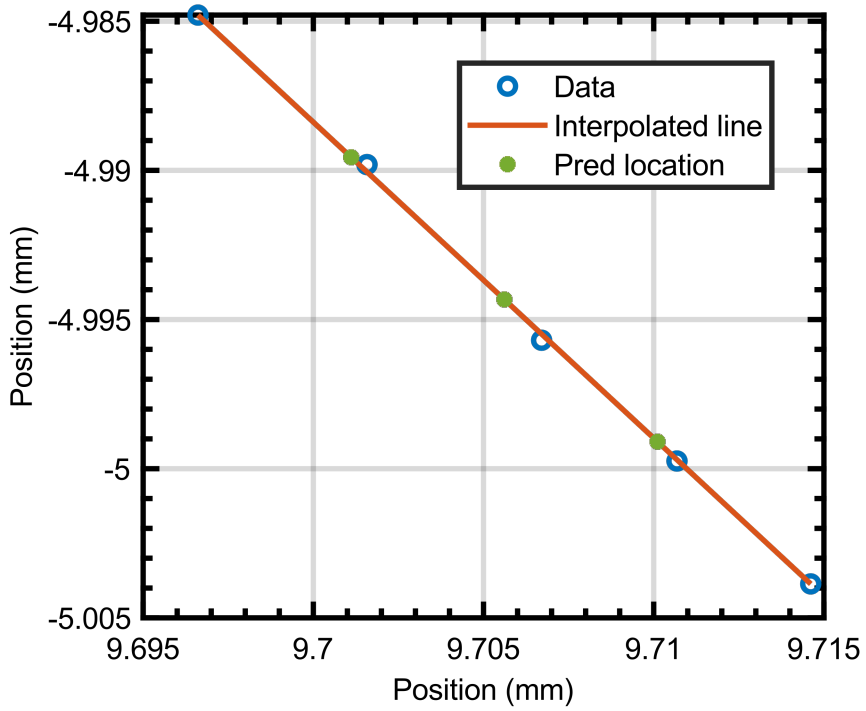


FIGURE 5.8: Sliding window (Fig. 5.7) of 5 points in a probe trajectory. End points interpolated and interpolated line marked at 1/4,1/2, and 3/4 distance. Line markings indicate predicted probe position using interpolation. Distance from line reported as EucD (euclidean distance), uniformity verified by (MD) mean deviation from uniform motion, distance from expected locations given by combined error CE, Pythagorean distance between expected and actual points.

This predicts improvement in alignment in SIM reconstructions in all cases. Using a translation stage with a mean interframe shift of 77.0 and 19.3 μm the errors in predicted location (ExpD) are respectively 3.2% (2.5 μm) and 6.1% (1.2 μm).

In freehand with an inter-frame shift of 77.6 and 8.6 μm the errors are respectively 32.9% (22.7 μm) and 21.2% (1.4 μm). Error is calculated in each five-frame set and averaged over all sets. While freehand motion does suffer from higher error, a higher frame rate vastly reduces the penalty paid (1.4 μm error is below the resolution of the bundle). A point that has not been previously raised is that square wave modulation may be more suitable for endomicroscopy. The work done in the SIM section showed that lines are robust to a constant phase alignment error of 18.3% between frames. This is due to the discrete nature of the lines and the fact that when considering a single pixel with intensity values $I_1, I_2, I_3 = 1, 1, 0$ an in-focus reconstruction of 1 is generated, which is exactly equivalent to $I_1, I_2, I_3 = 1, 0, 0$. Looking at any arbitrary individual pixel during constant misalignment, when misalignment $< 18.3\%$ is introduced, what happens is just a binary flip of 1 in $I_1, I_2, I_3 = 1, 1, 0$ in some pixels and the exact opposite flip of a 0 in $I_1, I_2, I_3 = 1, 0, 0$ in other pixels. Neglecting the defocus artefacts, which are non-negligible, it is worth considering this advantage that square wave modulation offers.

5.3.2 Line-pair width and robustness to misalignment

Freehand motion often introduces additional misalignment. When moving the probe, the square waves are aligned with the direction of motion, but since this is a feedback-driven approach, the direction recorded from a previous frame sequence is used, meaning direction does not match exactly, and adding complication the DMD has storage limits which means it only supports patterns in discrete rotation increments of 5.625°. By increasing the size of the lines, it's possible to

reduce the effect of misalignment by decreasing the relative misalignment. When using square wave modulation, this can have a secondary positive effect of bringing the relative misalignment within the 18.3% misalignment tolerance that lines provide.

The ability of the lines, with a varying LPW, to generate reconstructions which are robust to slight misalignment was characterised via numerical simulation of SIM reconstructions with a shift between raw frames. In the simulation, a circle representing a probe in contact with a uniform in-focus sample was superimposed with a line pattern. Three raw SIM frames were then acquired with the required phase shifts, with the probe also translated at some angle relative to the lines. The MAD was then calculated between the moving SIM reconstruction and an ideal reconstruction in an overlapping area of the bundle. A 5000×5000-pixel grid was used, the modulation depth was assumed to be 1 (i.e. ideal) and pixelation effects from the fibre bundle and discretization effects from the DMD were ignored. The circle's diameter was selected to match the fibre bundle (707 μm) and interframe shift (50 μm) used in the lens paper and bovine tissue studies below. The results showing the MAD as a function of line-pair width (LPW) and angular alignment error are shown in Fig. 5.9. Balancing the trade off between axial response and sensitivity to misalignment is application dependent.

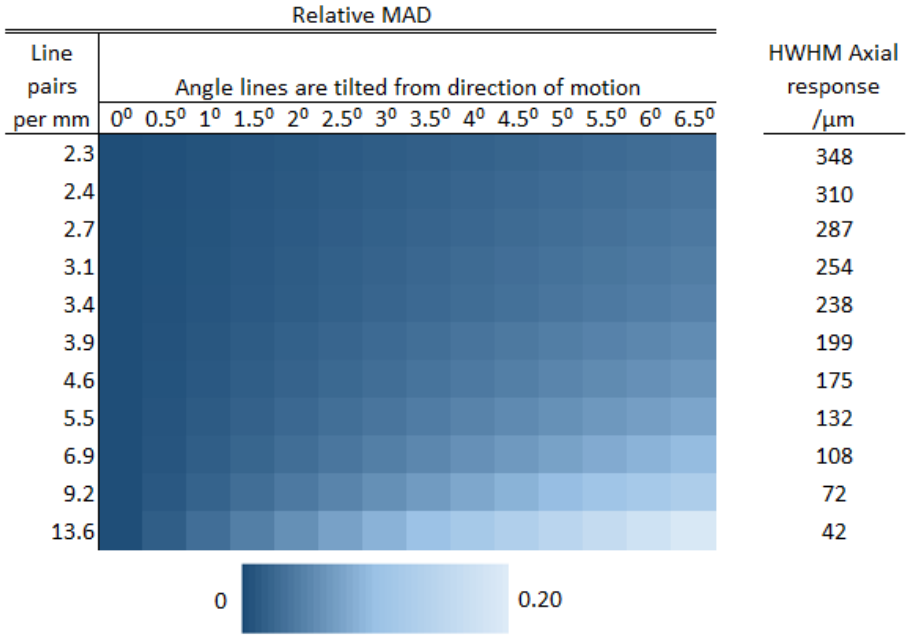


FIGURE 5.9: Simulation of SIM reconstructions when probe translation is not exactly parallel to the line orientation. The angle is measured between the direction of movement and line orientation, the plot shows relative MAD (In % error) of reconstructions with different LPWs. Simulation assumed a 707 µm probe diameter and a 50 µm shift between frames.

5.3.3 Automatic correction

Now that the linearity and uniformity assumptions have been verified, and using the LPW, previously characterised in terms of axial response and sensitivity to misalignment, the feedback mechanism is developed.

The LPW is selected to be 6.9 line-pairs per millimeter (lp/mm) or 145 µm. Axial response profiles show that with lines at 6.9 lp/mm out-of-focus light emanating from greater than 108 mm from focus is attenuated. The simulation using 6.9 lp/mm showed 95% similarity to static reference images (MAD=0.05) with an interframe shift of 50 µm and line-pair orientation off-parallel by 3°. At this LPW, pixelation artefacts are also minimal since the ~30 pixels per LPW is much greater than the minimum 6 DMD pixels (4.5 µm) required for SIM, it is also much

greater than the intercore spacing of the fibre bundle (3 μm). For completeness, the modulation depth was estimated experimentally to be 0.42 using the method reported by Hagen [134].

5.3.3.1 Program flow

The steps are broken down into three subsections: Image acquisition, template matching, orientation correction, and image reconstruction. They are all controlled by a custom-built LabVIEW (National Instruments) Virtual Instrument (VI). A summary of the programme control loop, with each component discussed in more detail below, is: (i) The Labview programme sends the required pattern angle to a micro-controller via the DAQ card. (ii) The micro-controller selects the pattern orientation and triggers the necessary patterns on the DMD. (iii) A 4 frame image sequence is acquired (Fig. 5.4) and a median filter is applied to the images. (iv) The angle and magnitude of motion between widefield illumination frames are determined by image registration using template matching. The shift is used for reconstruction and mosaicking, and the angle is used for the next update of the pattern orientation. (v) The raw frames are shifted to correct for motion between the acquisition of each frame and a SIM image is reconstructed, finally, each (vi) reconstructed image is added to a mosaic.

5.3.3.2 Image Acquisition

Identical to fixed motion compensation, using the VI, the 4 frame sequence is acquired using LabVIEW image acquisition libraries (*IMAQdx*). This consists of three SIM raw images, and one widefield (i.e. unstructured) illumination image

(Fig. 5.4). Once images are captured, the honeycomb structure from the fibre bundle needs to be removed. There are several ways to accomplish this including 2D Gaussian filtering[78], spatial averaging[150], and core localisation and interpolation methods[149]. The raw images are filtered using LabVIEW's *IMAQ Nth Order* filter as a median filter with a 4x4 pixel neighborhood (3 μm x 3 μm) — pixel values are arranged by intensity in descending order and the central intensity value is selected. More details, including a comparative table of previously reported methods, is included in Chap. 3.

5.3.3.3 Template matching

Template matching for both mechanical and automatic compensation is done by using an approach previously reported by Lewis for fast template matching [130]. First, a 250 × 250 pixels (187.5 μm x 187.5 μm) square is extracted from the center of the current wide-field illumination frame as a template, then calculate the normalized cross-correlation (NCC) matrix against the previous widefield illumination frame of 958 x 958 pixels (718.5 μm x 718.5 μm). The estimated shift is then taken to be the location of the maximum peak in the NCC matrix. Since the pulse structure (Fig. 5.10) is designed to give uniformly-spaced camera trigger pulses, then uniform motion is assumed between each raw frame over the acquisition sequence and estimate the interframe shift between the raw SIM frames as 1/4 the shift between the widefield illumination frames. Finally, the first and third SIM raw images are shifted relative to the second and generate a SIM reconstruction.

5.3.3.4 Pattern orientation correction and image reconstruction

As previously discussed, simply registering and shifting the images is insufficient to prevent artefacts since the line patterns are no longer correctly in phase. To reduce or eliminate the artefacts, the lines need to be oriented parallel to the motion of the probe.

The DMD can be used in several ways, a PC can output patterns to the DMD via an HDMI interface, or patterns can be preloaded onboard. Onboard pattern storage was used to take advantage of the faster pattern switching possible with this approach, and to avoid software lag when sending patterns over a video stream. Patterns of different angles (0-180°) are then loaded onto the DMD. The onboard storage supports storing 96 patterns, meaning 31 angles are supported (allowing also for a widefield frame), permitting a set of 3 patterns every 5.625 °.

Patterns are selected by the microcontroller using the angle determined in template matching. While the DMD does not allow us to select individual patterns at the hardware level, it does allow us to 'skip' patterns after a minimum display time. As shown graphically in Fig. 5.10 the total acquisition time of a frame is equal to the exposure time (t_{exp}) added to the product of the minimum display time (t_{min}) and the maximum number of unwanted frames between SIM patterns ($n_{skipped}$). When fewer patterns are skipped delays need to be introduced to keep the camera's frame rate constant. The formula for the theoretical maximum raw frame rate (fps_{raw}) is then given by:

$$fps_{raw} = \frac{1}{(t_{exp} + n_{skipped} \times t_{min})}. \tag{5.1}$$

The reconstructed frame rate is then 1/4 of the raw frame rate. With 30 skipped frames, a t_{min} of 325 μs, and exposure of 1,000 μs this supports a theoretical maximum raw frame rate of 93 fps or 23.25 fps per reconstructed image. This exposure isn't ideal for many applications, including the acriflavine stained bovine stomach tissue issue images reported below, where an exposure of 25,000 μs was used. The theoretical maximum frame rate for this exposure duration is 28 frames/s raw frame rate or 7 frames/s per reconstructed image. For consistency, the system was operated at 10 frames/s or 2.5 frames/s reconstructed throughout all experiments reported here.

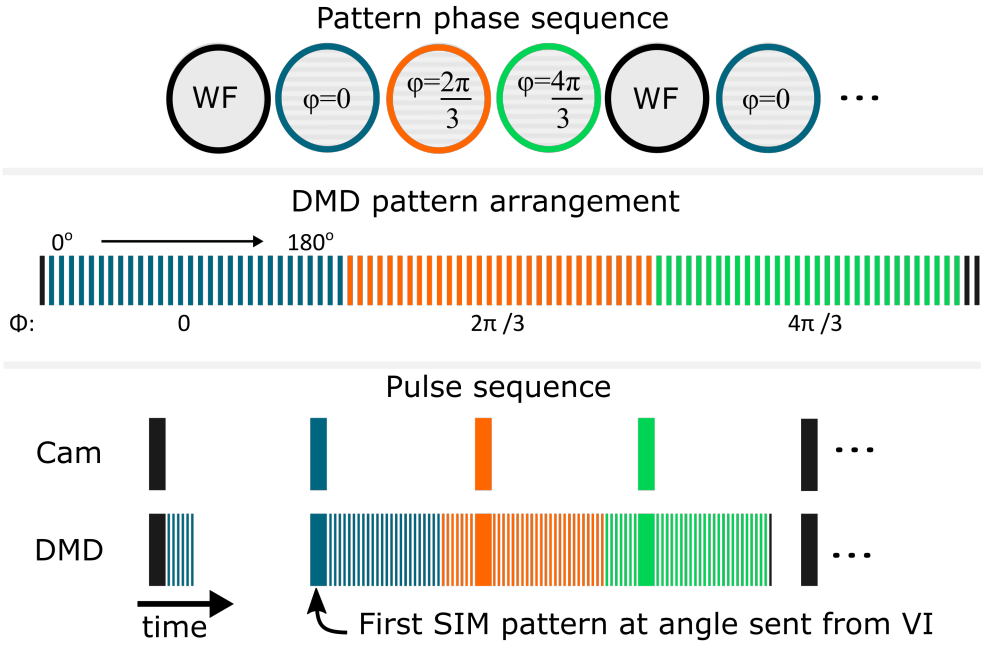


FIGURE 5.10: Pattern trigger structure. The top pane shows the phase sequence of patterns loaded onto the DMD. The second pane shows the storage of patterns at each angle by phase, the final pane shows the pulse sequence. Each pattern must be triggered in order, but the exposure can be varied. Non-uniform pulses are sent to the DMD, with longer exposures for the required patterns, synchronised with camera exposure trigger pulses, as shown in the lower pane. The pattern is designed so that the camera receives a uniform pulse structure.

During template matching, the pattern angle, in degrees, is scaled by 180, to get a voltage between 0 and 1, this is then sent as an analog voltage using

a DAC (National Instruments USB-6008) to a Pyboard. The Pyboard converts the received voltage back to the angle and, using the pulse structure shown in Fig. 5.10, selects the appropriate angle of illumination pattern and outputs the trigger pulses to the camera and DMD. During unwanted patterns, the Pyboard triggers the DMD every 325 μs and does not trigger the camera. To ensure the 4 frames are equally spaced in time the camera must be triggered uniformly and additional delays are introduced in DMD triggering to achieve this.

Once a sequence has completed, and the VI has received the final widefield illumination image in a sequence, it takes 135 ms for template matching and the required angle to be received by the Pyboard. At a camera frame rate below 20 frames/s, this voltage update is received by the Pyboard during the next cycle and the pattern orientation is changed after that cycle has completed.

5.3.4 Mosaicking

In order to assemble a mosaic as the probe moves, the position of each reconstructed frame relative to the previous is determined from the registration between widefield images. The SIM reconstructed image is cropped to a user-selected circle and added dead-leaf into the mosaic, over-writing any previous pixel values. The intent of mosaicking with motion compensation was to establish that, with compensation, it is feasible. More complex mosaicking is needed to correct applications where the probe motion is not approximately rigid translation. There is additional discussion in Chap. 3.

5.4 Results and Discussion

To demonstrate the improvement of SIM endomicroscopy over widefield the probe was first mounted to be stationary and in contact with a piece of lens paper stained with a yellow highlighter. To add a fluorescent background signal, the lens paper was placed on top of a piece of paper, also stained with a highlighter, and the two were separated by a thin plastic sheet, with approximate thickness 200 μm. This can be seen in Fig. 5.11. Fig. 5.12 shows the effect of optical sectioning and the improvement in contrast with a stationary probe.

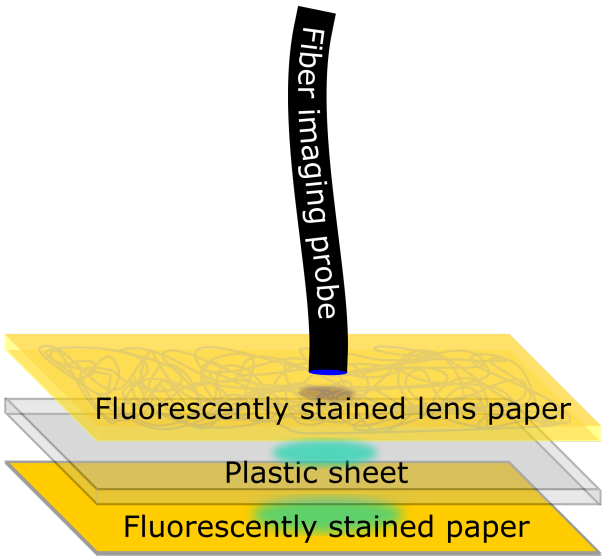


FIGURE 5.11: Lens paper imaging with fluorescent background.

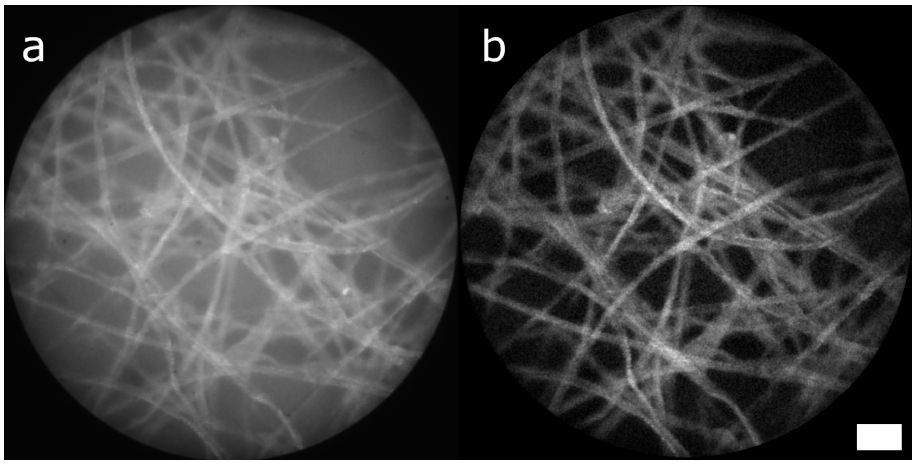


FIGURE 5.12: Fluorescently stained lens paper separated from a fluorescent background by a plastic sheet acquired with a stationary probe. Comparison of (a) widefield and (b) SIM images acquired with a fibre bundle probe. Bundle diameter: 707 μm. LPW: 6.4 lines per mm. Scale bar 100 μm

To validate motion compensation with a moving probe, lens paper and bovine stomach were prepared for imaging. The lens paper was stained with a yellow highlighter, and the bovine tissue with acriflavine hydrochloride using previously reported methods (Ref. [30, 151]). The probe was placed in light contact with the sample, then translated across the sample at a constant velocity that yielded a shift of 50 μm between each raw image, that meant the inter-frame shift was around 7% of the bundle's diameter. This inter-frame shift, and hence velocity, was chosen strategically based on the LPW. With 6.9 lp/mm it is an example of a case of greatest degradation due to motion artefacts (certain velocities lead to less degradation if the inter-frame shift happens to be close to an integer multiple of the LPW). This led to a shift between the two widefield frames, used for template matching, of 200 μm, this was a sufficiently large overlapping area for image registration. Examples of individual images are shown above the generated mosaics in Fig. 5.13 and Fig. 5.14.

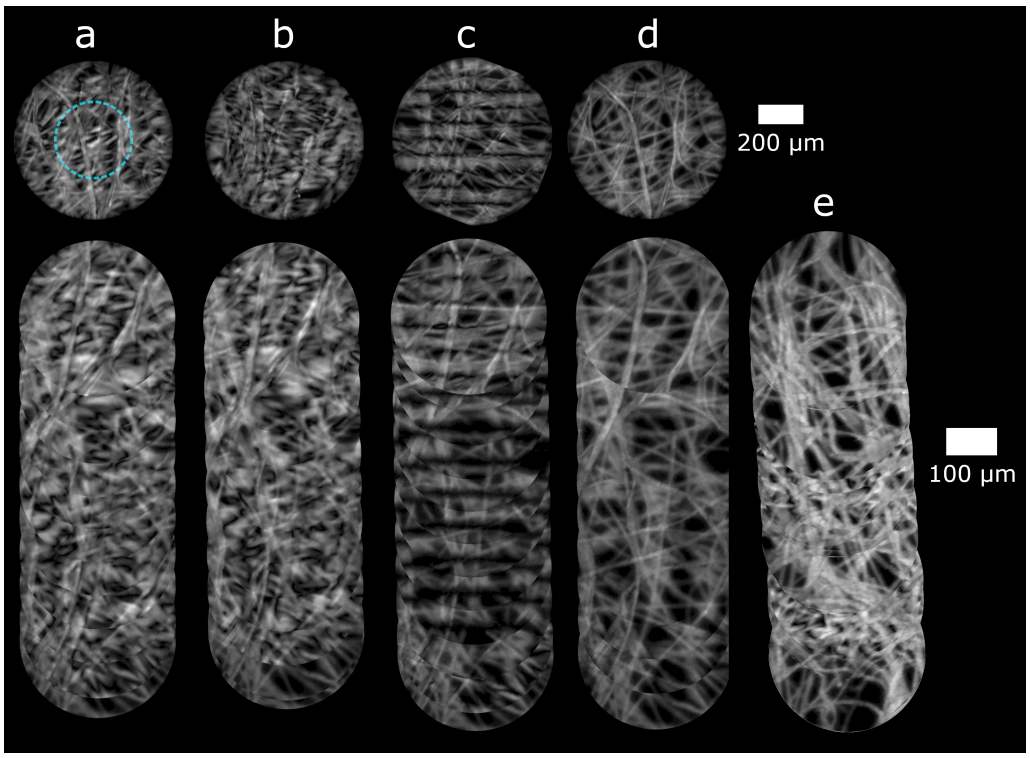
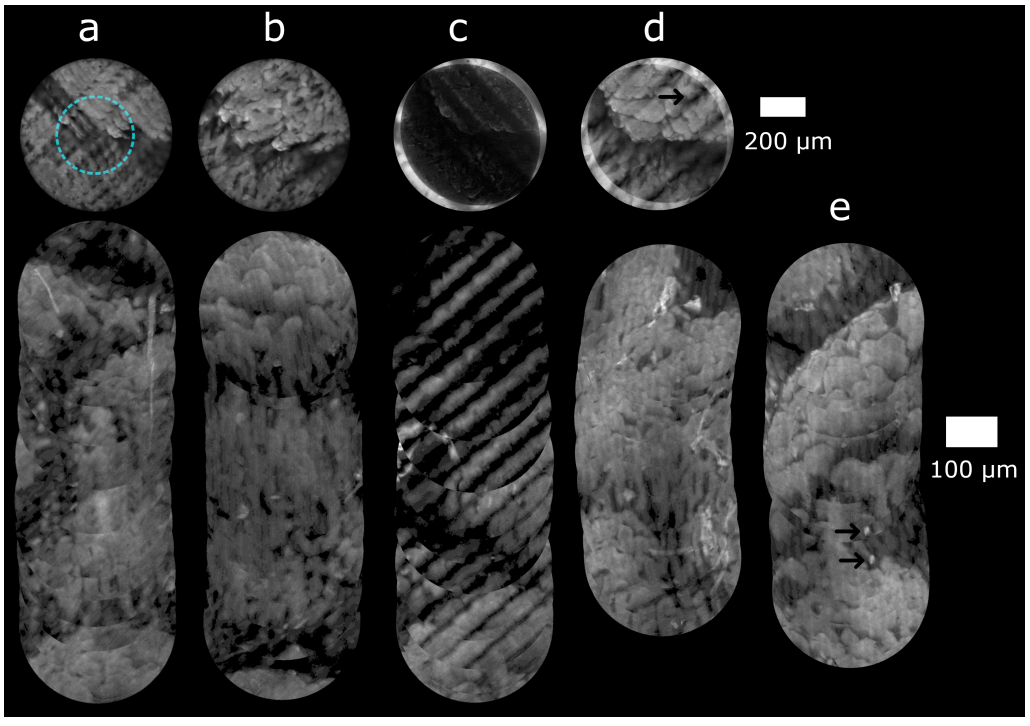


FIGURE 5.13: Moving probe SIM reconstructions and mosaics of lens paper stained with fluorescent highlighter. Lines at 6.9 lp/mm, frame rate of 2.5 frames/s per reconstructed image. Images a-d acquired with a translation stage (Velocity was 0.25 mm/s). a) pattern correction off / image registration off, circle shows cropped area for mosaicking, b) pattern correction on / image registration off, c) pattern correction off / image registration on, d) pattern correction on / image registration on, e) freehand images – pattern correction on / image registration on. Cropped area used for mosaicking to eliminate residual lines near edge of reconstructions.



	P0R0	P1R0	P0R1	P1R1
Lens paper				
NCC	0.44	0.39	0.97	0.97
MAD	0.85	0.91	0.14	0.13
Bovine stomach				
NCC	0.56	0.66	0.54	0.94
MAD	0.27	0.26	0.75	0.12

FIGURE 5.15: Relative MAD measurements as well as maximum of normalized cross correlation matrix for lens paper and bovine stomach tissue. Obtained for lines at 6.9 lp/mm and with translation of 50 μm between raw images. P. - Pattern orientation correction, R. - image registration. Number after P and R indicates on (1) or off (0).

FIGURE 5.14: Moving probe SIM reconstructions and mosaics of bovine stomach labelled with acriflavine. Lines at 6.9 lp/mm, frame rate of 2.5 frames/s per reconstructed image. Images a-d acquired with a translation stage (Velocity was 0.25 mm/s). a) pattern correction off / image registration off, circle shows cropped area for mosaicking, b) pattern correction on / image registration off, c) pattern correction off / image registration on, d) pattern correction on / image registration on, e) freehand images – pattern correction on / image registration on. Intensity values uniformly scaled by 2.5. Cropped area used for mosaicking to eliminate residual lines near edge of reconstructions.

Quantitative improvement can be seen by comparing individual reconstructions to a stationary reference. This is done by comparing the maximum of the normalized cross-correlation (NCC) as well as the relative mean of the absolute differences (MAD) for the images in Fig. 5.13 and Fig. 5.14. Four cases are compared for the 4 possible states of the system (pattern orientation correction on/off and image registration on/off). The values are reported in Tab. 5.15.

It is clear from Fig. 5.15 that introducing the motion compensation approach significantly reduced motion artefacts in both the lens paper and bovine stomach reconstructions. The improvement in the MAD value for lens paper was a decrease from 0.85 to 0.13%, and there was a decrease from 0.27 to 0.12% for the bovine stomach tissue. NCC scores increased from 0.44 to 0.97% for lens paper

and from 0.56 to 0.94% for bovine stomach tissue. The mosaics also show optically sectioned images with reduced motion artefacts. As discussed below, residual artefacts did appear in the individual images, however, cropping before mosaicking was sufficient to ensure mosaicking succeeded. In both cases, the combination of registration and pattern orientation correction yielded the greatest improvement over uncorrected reconstructions and reconstructions where only pattern orientation or registration was active.

Types of motion which can't be corrected by this technique include motion which occurs during the exposure. Artefacts from this occur, as they do in SIM endomicroscopy, widefield endomicroscopy, and confocal endomicroscopy. In SIM endomicroscopy and widefield endomicroscopy, they appear as blur, and in confocal, they appear as distortion. When features in a sample shift as the probe glides across they may appear as repeated features as in Fig. 5.14, which as can be seen did not break the continuity of the mosaic.

Some artefacts towards the outsides of the images were observed. These are seen as a more pronounced residual grid pattern in the image periphery, as marked with an arrow in Fig. 5.14 (d). The effect arises because, in a single widefield image, the intensity profile from out-of-focus depths is roughly uniform in the center of the probe but drops off near the edges. This is due to the effect of the edge of the bundle; out-of-focus depths nearer the edge receive less illumination than more central points, this was shown in the simulation in the SIM chapter, and is reproduced in Fig. 5.16.

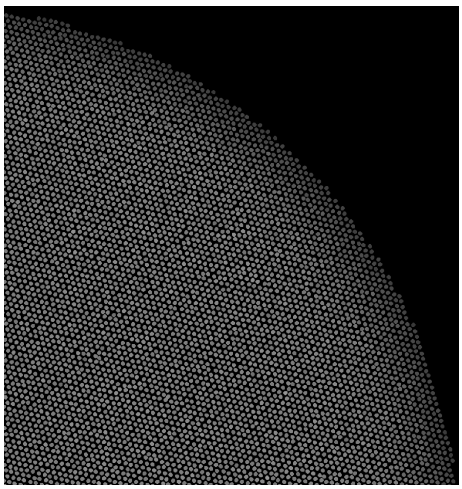


FIGURE 5.16: Edge artefacts with defocus. Z depth = 250 μm.

This does not matter for conventional SIM imaging as this is common to all three images, and the out-of-focus light is subsequently removed. However, it introduces a complication when reconstructing SIM images of a moving probe, as the three registered and shifted SIM images now have slightly different intensity profiles for the out-of-focus light contributions, leading to an incomplete removal. These artefacts can be mitigated at the expense of a smaller field of view by cropping the raw images to the overlapping region with roughly uniform intensity. This has a similar effect to introducing compensating illumination surrounding a smaller diameter bundle.

5.4.1 Edge effects correction via field normalisation

A way that has been previously reported to reduce edge effects is via field normalisation. This is accomplished by taking a flat field estimation and divide each image by the estimation. This is a sample dependent correction since the scattering properties change. For this demonstration lens paper labelled with fluorescent highlighter separated from a fluorescent paper by a ~ 200 μm plastic sheet is

used. To get a flat field estimation, the probe is rapidly translated over the tissue 117 images are acquired, the intensity of these images is then averaged to get a flat field estimation (FFE). Next the inverse of the FFE is used to generate a bias, and this is multiplied by each raw image to get a flat field corrected image (FFC). Fig. 5.17 shows the improvement in feather near the edges. Following the same approach to gain a widefield FFE, the probe is then translated on a stage and get a SIM FFE, averaging over 51 motion compensated SIM images. This further shows the reduction in edge effects.

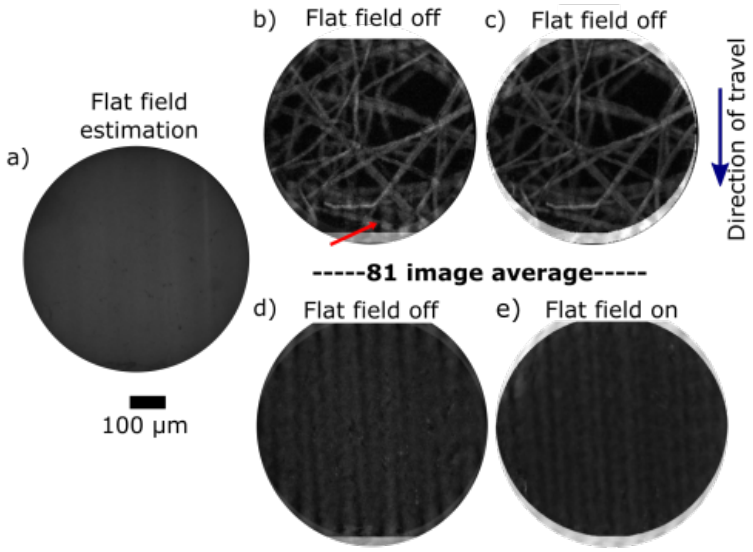


FIGURE 5.17: Probe translated across lens paper seperated from a fluorescent background, (a) average of 117 widefield images, (b,c) individual reconstructions, (d,e) average of 81 motion compensated SIM images, with flat field correction on/off.

5.4.2 Defocused line artefacts

Streaks in the final mosaic occur when the motion of the probe is not perfectly parallel to the alignment of the patterns, as can be seen in Fig. 5.13 (b,d,e). The contribution to streaking attributable to misalignment artefacts arises due to the finite number of angles accessible, due to the limited number of images which can be

stored by the DMD. This means the direction of translation has to be rounded to the nearest supported angle. However misalignment may not be the only cause, these may also be defocus artefacts, the result of using square wave lines rather than pure sinusoidal illumination patterns, this was discussed in far more depth in SIM - defocus artefacts. The case that defocus artefacts contribute is supported by a few factors: the presence of these artefacts in multiple mosaics, and the fact that they appear along the entire mosaic. This isn't expected in freehand motion since alignment changes along the probes trajectory. These are, however, unlikely to be edge artefacts, the cropping circle was set to half of the probe's diameter, where the widefield images had roughly uniform intensity.

While the field-of-view of the individual motion-compensated moving SIM images is reduced, compared with conventional SIM there is no effective reduction of the usable field of view in any case. When the probe is stationary the device performs as conventional SIM; when the probe is in motion, conventional SIM does not work at all. With motion compensation, the field of view is reduced and mosaicing is then possible, allowing for a larger effective field of view.

5.4.3 Other artefacts

Artefacts not previously mentioned include repeated features in the freehand mosaics for the bovine stomach. These were not observed in the tissue paper study, and so a likely cause is misregistration of tissue deformed during translation of the probe in contact mode (only rigid transformations were allowed for by the algorithm). Additionally, the latency associated with pattern selection meant that

pattern orientation was always one acquisition sequence behind, leading to artefacts when the direction of travel changed between acquisition sequences. This is mostly caused by software limitations. While template matching was benchmarked faster than the maximum camera frame rate (60 Hz), the updated angle was still received by the pyboard shortly after the next sequence had started. Another source of latency was the speed at which the LabVIEW VI sent the angle to the Pyboard via the DAQ. This could be reduced by using a serial port to communicate with the Pyboard directly instead of analog pins. A topic for further investigation is devising a scheme that would take advantage of both the fast template matching and serial data port to eliminate the pattern orientation lag.

With mounted operation using a translation stage, in a configuration where angles don't need to be updated each cycle, artefacts caused by latency are mostly eliminated since the angle is the same for a large number of acquisition sequences. Conversely, handheld operation is more susceptible to artefacts resulting from deviations in a movement parallel to line orientation. These artefacts would be reduced with a faster camera and DMD pattern selection time.

The frame rate of 2.5 frames/s was due to the lower duty cycle at higher frame rates forcing the camera exposure to be reduced. This arises from the use of a low-cost off-the-shelf DMD with an onboard LED, which without adding active cooling, had optical power limits, and a DMD controller which does not allow patterns to be selected arbitrarily. Unfortunately, as each pattern has a minimum display time of 325 μs, as the frame rate is increased the duty cycle decreases due to an increased fraction of the overall time needed to wait while unwanted frames

are skipped. Pattern selection time and optical power are not fundamental limitations of this approach. Pattern selection time could be increased with a more sophisticated DMD controller which allowed the selection of arbitrary patterns. The power at the fibre tip was measured to be 95 μW, and the fibre coupling losses were measured to be 89%, power could be increased by adding active cooling to the DMD, optimising lost light at the objective/bundle interface, or exchanging the DMD's built-in lens and LED with a high power light source.

5.5 Translation to HiLo

At this point the work done in motion compensation is adapted to HiLo microscopy and show improvement is realised there as well. The primary intention of this section is not to demonstrate an optimal HiLo system, only to demonstrate that our motion compensation approach can be easily adapted to HiLo illumination.

5.5.1 Introduction

HiLo illumination, developed in 2012 by Ford, T., et al. [86], is a numerical method of optical sectioning that requires only two frames to reconstruct a single image. The first frame of the two frame sequence uses line patterns, of a tunable spatial frequency I_s, where "s" indicates structured illumination, projected onto the proximal bundle face, and the second frame with uniform illumination I_u. It is based on the same idea as SIM: modulate in-focus information, locate the modulation, and finally, reconstruct an image with the modulation removed.

Since one of the frames is acquired with wide-field (unpatterned illumination), it ought to be possible to compensate for motion with no modifications to either the pattern orientation or additional frames in the sequence.

The HiLo principle is this: when projecting line patterns onto a sample which consists of both in-focus and out-of-focus components, it's possible to determine where local variations in contrast are (i.e. located the in-focus information) and produce a contrast map of values with high contrast regions having greater modulation than low contrast regions. If a low pass filter is performed on this map, an image of just the in-focus low-frequency features is obtained, including the uniform plane, but by the nature of the process to generate the map it excludes high-frequency features. These components can be included by adding an image (I_u) filtered with a high pass filter. This is possible because the high pass contains necessarily in-focus features, since diffraction inherently blurs out-of-focus features if out-of-focus, they are attenuated.

5.5.2 Original system description

Ford, T., et al. reported the optical system for HiLo microscopy shown in Fig. 5.18. Light from a laser is guided to a galvanometer and directed between two paths corresponding to the I_u and I_s frames. The first path, is through a Ronchi ruling (50 lines per mm) which generates the lines. In the second path, light is guided from the source and simply reflected off a mirror and directed onto the bundle face. The half-wave plate is used to polarise light on the Ronchi ruling arm, and the polarising beam splitter is used to reflect unpolarised light, and transmit the polarized. The light is then incident on the bundle, and fluorescence returns

through the objective and is imaged onto a camera, via the dichroic beam splitter, tube lens, and emission filter.

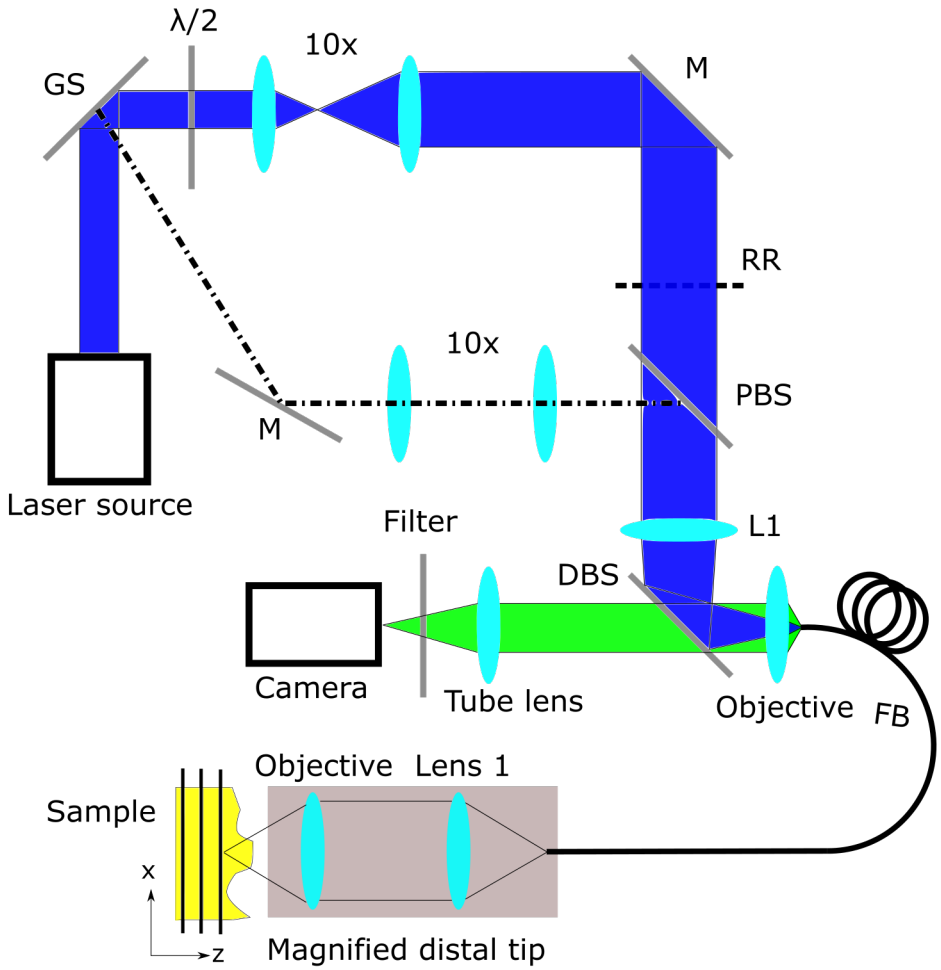


FIGURE 5.18: HiLo schematic, as reported by Ford, T. et al. GS - galvanometer scanner, $\lambda/2$ - 1/2 wave-plate, M - Mirror, RR - ronchi ruling, PBS - polarising beamsplitter.

For the study reported here, the same system was used as for the SIM studies. There are a few key differences from the approach reported by Ford, T. et al., the SIM system uses incoherent illumination and a DMD to generate the modulation, making it optically simpler since the galvanometer is no longer needed. Nor is the PBS or half-wave plate needed since they were only used to minimise the losses when using a non-polarising beamsplitter. A lensless fibre is also used, which does not fundamentally change the sampling of the bundle.

5.5.3 HiLo processing algorithm

I used a Gaussian kernel to preprocess images and remove the pixelation structure. Unlike SIM, where processing does not require frequency domain operations, HiLo mainly uses the Fourier domain for processing. Consequently, the high spatial frequency cladding structure **must** be completely removed before any HiLo processing. The highest frequency the bundle will transmit is approximately given by the Fourier transform of the intercore spacing. As will be discussed the cut-on value of high pass filter, which contains approximately half of the intensity of the reconstructed image, will always be below this, this means, if not suppressed, the cladding will always be reconstructed, and by the nature of the reconstruction scheme, amplified. Fibre bundles consist of pseudohexagonal structures. This makes them difficult to process with spatial frequency domain approaches. The reason for this is, using an approach like disk filtering one may inadvertently filter features with dimensions around the intercore spacing. There are fibre bundles manufactured with perfectly regular core spacing which are more amenable, however, quartz bundles are distinctly non-periodic. For this reason, it is necessary to remove the fibre bundle pixelation structure before doing any processing. Simple Gaussian blurring can be used however the image kernel must be sufficiently wide to attenuate the core pattern and consequently, leads to a loss in resolution. The cladding removal section of this thesis discusses how Ford, T. et al. chose to handle this in Chap. 3.

Once the cladding is suppressed, the images can be processed. Fort, T. et al. first treat the uniform and structured images as consisting of in and out-of-focus components and describe how the out-of-focus components are removed in both.

The uniform component is given by:

$$I_u \quad = \quad I_{in} + I_{out} \tag{5.2}$$

and the structured image is given by:

$$I_s \quad = \quad I_{in}(1 + m\sin(2\pi\kappa_g x)) + I_{out} \tag{5.3}$$

where, x is the position along the modulation, m is the modulation depth, and κ_g is the modulation frequency. The next step is to locate regions where the grid pattern appears. This is done by taking the ratio of the structured image to a Gaussian filtered structured image to normalize the global image in a way that corrects for local intensity variation. The spatial frequency cutoff of the Gaussian low pass filter must be much less than the frequency of the lines. This yields a formula for contrast:

$$C \quad = \quad \frac{I_s}{< I_s >} - \frac{I_u}{< I_u >} \quad \approx \quad m\sin(2\pi\kappa_g x) \tag{5.4}$$

the next step is to take a low pass filter of the product of the contrast map and I_u.

$$I_{low} = LP[C \times I_u] \tag{5.5}$$

this yields an image where the in-focus low spatial frequency components have

been selected and line modulation is blurred. A high pass filter of the unstructured image is then taken. The final reconstructed image is given by:

$$I_{HiLo} = I_{high} + \nu I_{low} \tag{5.6}$$

where ν is an adjustable parameter to account for the intensity loss during the Gaussian smoothing operation.

High frequency features, especially those near the frequency cutoff of an objective or a bundle in our case, are necessarily in focus, however, the power of HiLo lies in the ability of the contrast map to select in-focus low frequency features. Looking at the contrast map, the modulation will only appear in-focus and will oscillate around zero. In both the first and second expression, the denominators serve as both a local average reference (smoothed by a low pass filter with a frequency cutoff much greater than the modulation), and a way of normalizing the intensity locally. While the first term serves as a selection term it has the overlapping grid pattern imposed on top, the subtraction of the second term allows the modulation to oscillate around 0 instead of 1. Taking the low pass filter of the contrast map, I_u product yields an approximation of the in-focus region. Since high-frequency terms must be in-focus, and they were subtracted out of the contrast map, it's then clear why they are added back in to reconstruct the image. This formula also shows how unsuppressed cladding can be amplified.

5.5.4 Motion compensation

With HiLo there is no need to rotate the structured image in the direction of motion. This elimates the triggering protocols required for motion-compensated SIM. The reduction in artefacts with HiLo using a LPW of 31 μm is now demonstrated, selected since at this spatial frequency the bundle transmits the modulation with ~95% contrast. A probe was translated in freehand operation, and the motion compensation approach previously described was used. The following parameters were used, a low pass kernel width 6.4 cycles / μm. The high pass filter used was the inverse of a Gaussian kernel with standard deviation 0.8 cycles / μm. Modulation frequency κ = 5 μm. Fig. 5.19 shows contrast improvements with motion compensation over HiLo as well as a reduction in incorrectly subtracted light. The images are reported unnormalised, however, in order to meaningfully compare contrast in the intensity profiles, before taking profiles, the images were independently normalised to 1% saturation. There is a clear improvement when using HiLo over widefield illumination. Mosaics were also generated (Fig. 5.20), these show improvement in contrast over widefield images, and motion artefact improvement.

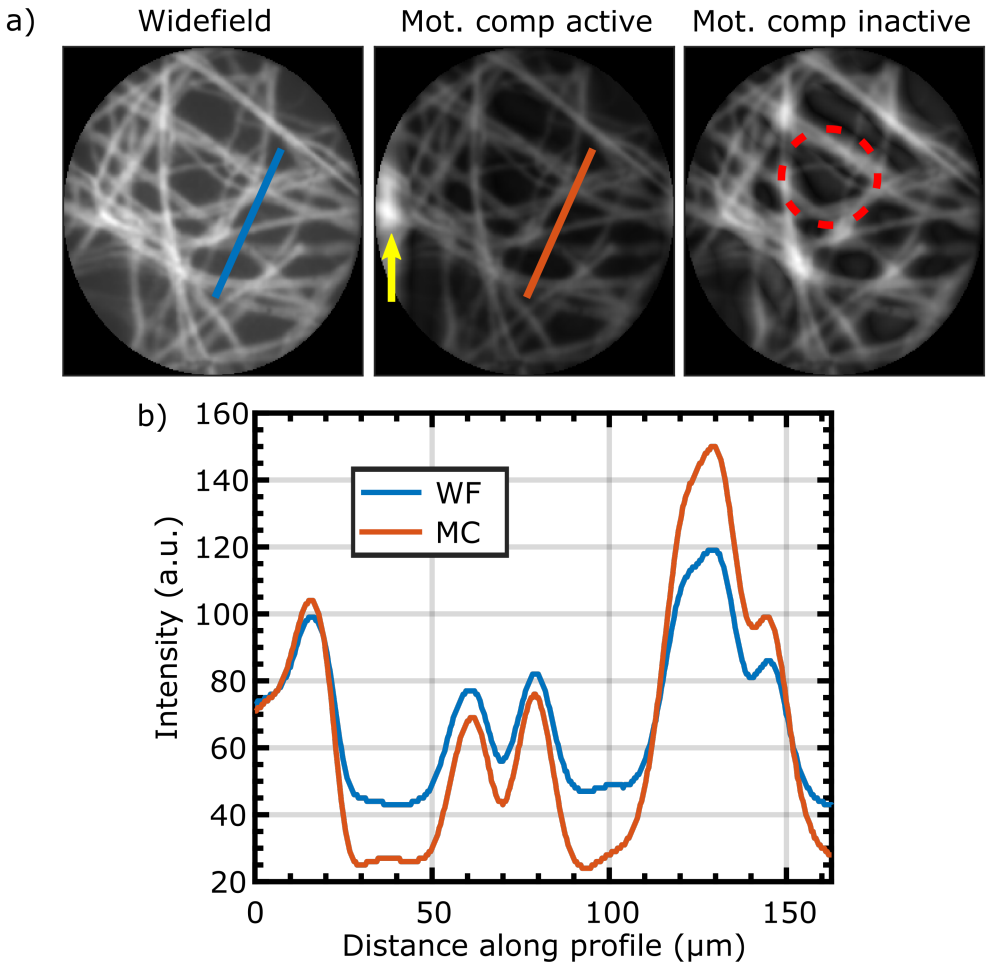


FIGURE 5.19: (a) Images of lens paper acquired with a moving probe. (b) Intensity profiles of motion compensation (MC) and widefield images. Arrow indicates artefacts from shifting probe, circle indicates incorrectly subtracted images due to motion. Widefield and motion compensation images independently normalised, inactive motion compensation image normalised by same factor as motion compensation active image.

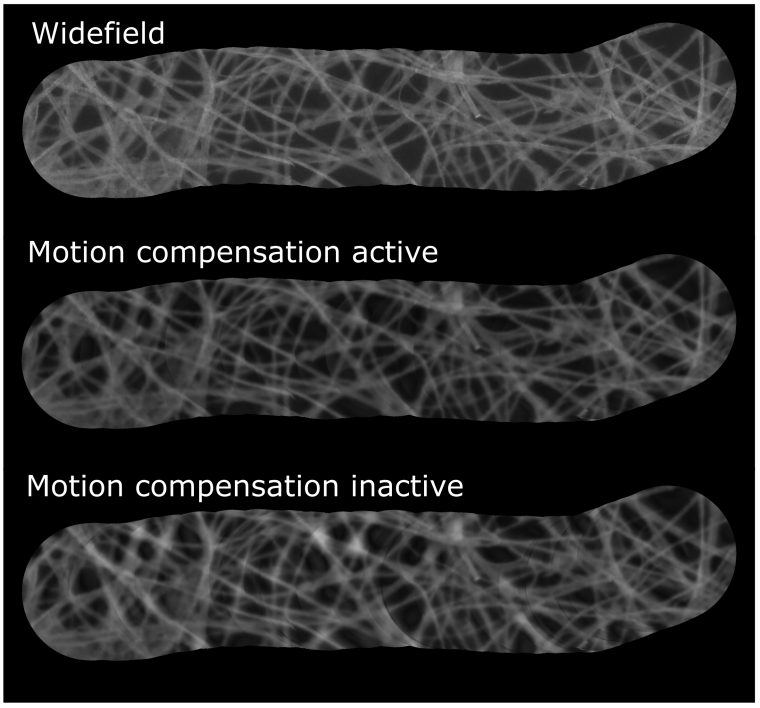


FIGURE 5.20: Mosaics generated with widefield illumination, motion compensation, and with motion compensation inactive.

It may seem appealing to reduce the frequency of the modulation as low as possible to reduce the penalty paid during the smoothing process of the contrast map. However, the extent of this smoothing is limited by the frequency of the modulation, and the frequency of the modulation is limited by the core spacing of the bundle. This means HiLo, like SIM, has a fundamental limit to the modulation frequency determined by the transfer function of the bundle.

5.5.4.1 Edge effects

Artefacts with HiLo are different from SIM, this is due to a few factors: limited motion between frames, and as will be discussed below, the low pass filter smooths features to beyond their actual location. However, edge artefacts in the high pass image will be likely unaffected. In the SIM section, it was shown that since the intensity is lower near the edges of the bundle, edge artefacts result

from reconstructions of the same feature in adark portion of one frame and in the light portion of another. The primary reason for the global intensity difference is that out-of-focus light preferentially returns to central cores. Since high frequency features are, by definition, in-focus it follws that the high pass image would not suffer from the same artefacts. The same can't be said for the low pass image.

One problem that's apparent with this approach is that when creating a contrast map, the structured image need to be blurred by a Gaussian low pass filter. It is necessary that the spatial frequency of the low pass filter be much lower than the modulation frequency. Furthermore, as developed in the SIM section, there are discretization effects due to the fibre bundle when sampling lines. These effects can be reduced when a low spatial frequency is chosen, this implies a large Gaussian kernel to blur the image.

To see how this leads to artefacts that can be seen in the spatial domain. Consider a Gaussian Kernel:

$$K = \frac{1}{16}\begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{pmatrix}.$$

A convolution with a Gaussian is the sweeping of this kernel over a 2D array of pixels, where the sum of the point-wise product of the kernel and a local region become the value of the central pixel. The problem with the bundle is there is a sharp intensity transition between the bundle, and the black background. Applying this convolution kernel near the edges of the bundle will always lower the intensity in the denominator. One could imagine the dot product of a uniformly

illuminated bundle with normalised intensity at a region near the upper left edge and kernel as looking like:

$$K * T_{image} = \frac{1}{16} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 4 & 2 \\ 1 & 2 & 1 \end{pmatrix}.$$

where T is a template, of size K, extracted from the image. This means the intensity at the edges is reduced by 3/16. For a small region, such as a 3x3 kernel with a 500x500 pixel image, the effect is small. However if the criteria is that a large kernel needs to be used, this can lead a substantial reduction in intensity around the edges.

Effectively the trade off is this: narrow lines lead to better modulation, fewer edge artefacts, and less blur when merging the contrast map and I_u, they also lead to pixelation artefacts with the bundle, which may contribute substantially to reconstruction error, and negate any benefits. Larger lines on the other hand, still filter light far from focus, but can adversely contribute edge artefacts.

5.6 Conclusion

Endomicroscopy is a useful tool for point-of-care diagnosis of epithelial cancers. A low-cost fibre bundle based SIM endomicroscope can be built with a probe area much less than the area of a conventional biopsy, motivating the need for mosaicking. During the three frames, SIM acquisition sequence motion between raw frames can generate artefacts which limit mosaicking. The findings presented here show that by re-orienting patterns to the direction of motion, and

by registering and shifting the images, it is possible to greatly reduce these artefacts. They further show the technique can be directly adapted to HiLo. This will further advance the clinical relevance of SIM-based endomicroscopy.

Chapter 6

Optical sectioning with alternative structured illumination patterns

6.1 Introduction

SIM with three-phase line modulation has limitations including the need for lower frequency modulation to ensure adequate contrast transfer through the bundle and the presence of residual line artefacts. This chapter presents *preliminary results* and conclusions to evaluate the feasibility of modulation using patterns other than lines, which have not yet been adapted to endomicroscopy.

The patterns that will be investigated will be generated on a pixel or superpixel (block of pixels) grid. Two pattern types will be investigated are Tetris, and Four Quad. Tetris is a novel implementation of random speckle illumination. Four Quad is an approach where in any four pixel (or superpixel) square region on the DMD, a pixel which is "on" rotates through each of the four quadrants with all others being "off".

2D modulation allows spatial frequencies three times higher than SIM with lines to be used. The minimum line pair width (in DMD pixels) required for SIM is 6. While in the 2D case an "on" pixel can be separated from another "on" by a single "off" meaning the spatial frequency is 3 times higher. There are also

advantages to bundle sampling. Consider sinusoidal modulation with spatial period 19.2 μm (6 cores) being sampled by a row of pixels. A bundle transmits perfect contrast if each core sampled only a single point. However, when a bundle samples the sine wave patterns, the cores do not sample a single point, each core averages the intensity across its face, this averaging operation, when compounded with the effects of phase shifting this leads to diminished contrast. This was manifest in Fig. 4.13(a,c), the study determined the reconstructed intensity as a function of line pair width. This gives a metric of modulation transmission that also includes the effects of phase shifting. As modulation frequency is lowered (lines are enlarged) the transmission percentage reaches an asymptotic plateau. While the value of the asymptote is clear, it is difficult to say at which modulation frequency the intensity has reached the maximum. To avoid any ambiguity the curve can be parameterised using the 50% peak value. With SIM, 50% intensity occurs at a modulation depth of ~ 40 μm, while with Tetris reconstructions, demonstrated below, 50% intensity occurs at ~ 10 μm. Reduced artefacts are also demonstrated by comparing the relative mean of the absolute deviations (rMAD). With a SIM modulation depth of 54 μm, and Tetris pixel size of 27 μm, the rMAD's are 32% and 28% respectively. Additionally a superpixel shifting technique for further rMAD improvement is demonstrated.

Four Quad, a different approach, demonstrates improved optical sectioning in simulation, and experiment with and without using a bundle. In the bundle free case, it was experimentally determined that Four Quad and SIM have a HWHM of axial intensity response to defocus of 12, and 18 μm. In the fibre bundle case, the respective HWHM's are 21 and 22 μm respectively. Problematically, in both

approaches, improved performance comes at the expense of artefacts, exhibiting a factor of 3 greater artefacts in the bundle free case (30% rMAD), and factor of 2 in the bundle case (20% rMAD). The reason for the artefacts is described in the half-pixel problem, and resolution of this would make way for improved sectioning.

6.2 Background

Random speckle illumination has been reported for optical sectioning [152–155]. First in 2006 by Ventalon, C. et al. and 2013 by Hoffman, Z. et al. sectioning are obtained by projecting high spatial frequency features onto focus and taking the differences between images to extract in-focus features [153]. Ventalon et al. reported modulation with coherent laser speckle, and Hoffman et al. imaging a ground glass plate, incoherently illuminated, on to focus. The speckle approach relies on speckle grains that possess short-ranged spatial intensity correlations which are determined by the average speckle grain shape, which in turn is dictated by the diffraction [156]. This work has also been expanded using an SLM [156] to introduce non-local correlations as well, and tuning the sequence of random patterns to change to resulting field and change the topologies and speckle statistics [157, 158]. For the fibre bundle approach, the spatial distribution of intensity requires near-perfect modulation, at least on the order of a single core's diameter. Otherwise, bundle sampling, both proximally and distal fluorescence leads to greatly diminished contrast (Fig. 6.7).

The subtraction based demodulation process reconstructs regions where modulation appears, and regions where the modulation has blurred, are excluded. A

schematic of the system presented by Hoffman, Z. et al. are shown in Fig. 6.1. Ground glass has a few properties that lines do not: first, the intensity is slower varying than the sharp edges of the lines, second, modulation is not just "on" or off, there are intermediate intensity values, and third, modulation is in two dimensions. Imaging is done by rotating the plate and capturing images at different rotations. A large number of images are required to reduce modulation artefacts in reconstructions.

The demodulation formula used by both authors takes the differences between sequential images, using [152, 153]:

$$I_{RS} = \frac{1}{2N} \sum_{n=1}^{N} \sqrt{(I_n - I_{n+1})^2} \tag{6.1}$$

where I_{RS} is the processed image using the random speckle method, I_n is a raw image in a sequence, and I_{n+1} is the image acquired a short time later. Hoffman et al. with ground glass images use 40 images. With a 120 Hz camera, this implies a 3 fps acquisition rate, assuming there was no latency introduced by processing. 40 patterns mean the rame rate is around an order of magnitude lower than three-phase SIM; this is a disadvantage for endomicroscopy.

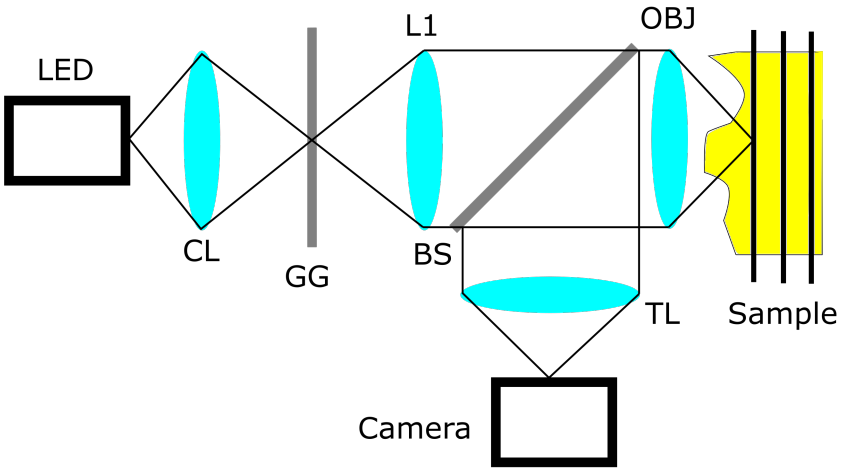


FIGURE 6.1: Random incoherent speckle illumination diagram. Light emitted from an LED is collimated by a lens (CL), which is incident on a rotating ground glass plate (GG), from there it is relayed onto a sample via a lens (L2), beamsplitter (BS), and objective (OBJ). The returning reflection or emission is then guided to a camera via the beamsplitter.

6.3 System description

All approaches described above use the same optical system (Fig. 6.2). Depending on the context the fibre bundle in the diagram may be removed and is typically replaced with either a mirror or a smooth metal plate. Light is emitted from the built-in LED of a DMD (TI DLP Lightcrafter 3000) and relayed through its built-in lens, then a 50 mm achromatic lens (Thorlabs AC254-050-A-ML), an emission filter (Thorlabs FES0450), off a dichroic beam splitter (Thorlabs DMLP490), onto a fibre bundle (Fujikura FIGH-30-650S), the returning fluorescence is guided through a finite conjugate (160 mm) objective (Plan 10/0.25), through the dichroic beam splitter and excitation filter (Thorlabs FEH0500), onto a camera (PointGrey FL3-U3-13S2M-CS). The fibre bundle has an inter-core spacing of around 3 μm.

DMD and camera pixels dimensions were determined at the object plane. 1 camera pixel = 9 μm, and 1 DMD pixel = 2.5 μm.

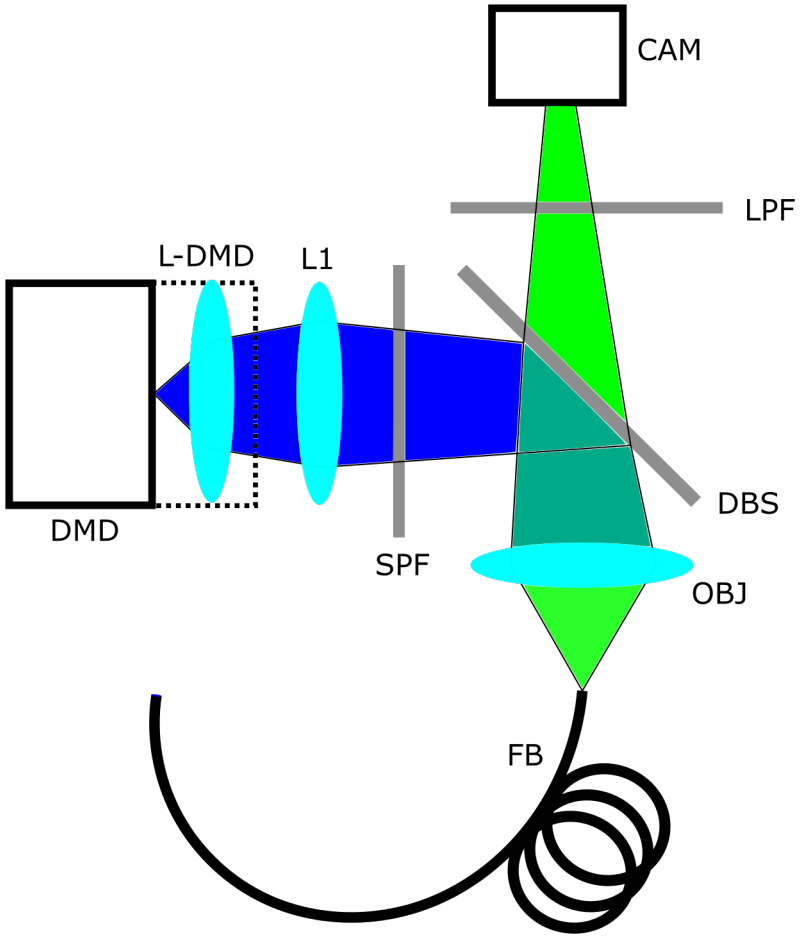


FIGURE 6.2: Optical diagram of a DMD based fluorescence endomicroscope. DMD - digital micromirror device, L-DMD - built in lens, L1 - compensating lens, SPF & LPF - Short, Long pass filter, OBJ - objective, FB - fibre bundle.

Two parameters were studied for all methods under investigation: intensity response to defocus and in-focus artefacts. Defocus intensity response is determined by moving a mirror through focus and determining the distance at which the signal has decayed to 50 %. In-focus artefacts are characterised using the mean of the absolute deviations relative to the mean (rMAD). A representative curve is shown in Fig. 6.3, shown is the intensity response and relative mean of

the absolute differences curve. The diagram identifies the HWHM value of the axial response as well as the in-focus artefacts rMAD.

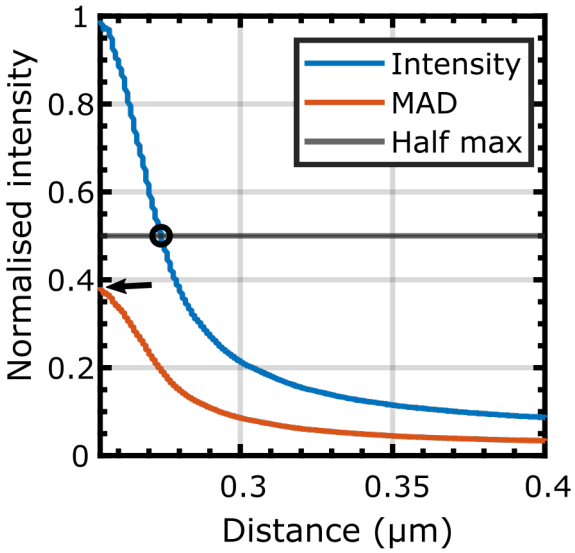


FIGURE 6.3: Intensity response and artefact characterisation of a Tetris system under defocus. N=4 patterns. Half-width half-max identified as circle. Absolute mean of the absolute differences is used to plot artefacts with defocus. Arrow identifies percent in-focus artefacts.

6.4 Tetris

Tetris illumination starts with the same basic idea as random incoherent speckle illumination, but instead of using ground glass, a DMD is used to generate random patterns. Patterns are used with pixels that are either 'on' or 'off' during the entire acquisition of a frame. DMD illumination is different from ground glass since there are no intermediate intensity values. Demodulation is done using the same formula as random speckle. When deriving random speckle illumination with ground glass, Hoffman dedicated a significant section to ensuring correlation between frames was at a minimum [153]. There is no need to repeat the same

steps since random pixel values are assigned digitally. When comparing two images for each pixel, there are four microstates; they are 11, 10, 01, 00. With 1 being "on" and 0 being 'off'. They are the same 50% of the time.

First, a feasibility simulation is conducted to evaluate whether any sectioning improvements are expected. The simulation considers the axial response to defocus of a thin fluorescent mirror moving away from defocus. Relative intensity is determined at the SIM half power-point for 40 frame Tetris. A SIM pattern is generated with an arbitrary spatial frequency of ν. A superpixel size equal to 1/2 the SIM line pair width is used. The Gaussian kernel width that yields a reconstructed intensity of 50% is determined by iterating through values until a threshold of 50% is satisfied. The raw patterns for each of the test methods are then by the same Gaussian kernel. The reconstructed intensity of the Tetris is 68% indicating inferior perforance with Tetris. At a spatial frequency 3 times higher that SIM, the simulation returns 51%, effectively putting it on par with SIM. Experimentally, comparable sectioning, signal improvements via improved modulation transfer, as well as reduced artefacts will be shown, although as will be seen this comes at the cost of number of frames.

Random patterns were generated, a 25 x 25 pixel square extracted from a single pattern is shown in Fig. 6.4. It looks similar to a familiar QR code. Matlab's random number generated is used to generate binary values for each pixel. Camera-based triggering is used with a camera operating at 120 Hz. After acquiring a frame, a trigger pulse is sent to the digital micromirror device to change patterns.

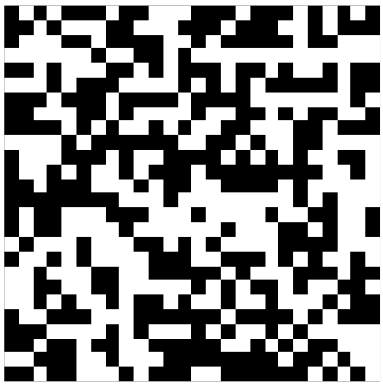


FIGURE 6.4: 25 × 25 pixels extracted from a random pattern used for the Tetris Method.

6.4.1 Reconstruction artefacts at focus

In random speckle imaging, Hoffman et al., [153] used 40 raw frames per reconstructed image. Since it is the number the authors reported as necessary for random speckle to reduced artefacts in images. To determine this, they computed the standard deviation in each pixel in the reconstructed image. The benefit of the DMD approach and known patterns is that this can first be simulated. The relative mean of the absolute differences as a function of the number of patterns in each reconstructed image is determined. The penalty paid for more faithful reconstructions is the number of patterns. The acquisition time is determined by (frame rate)/(number of patterns) this is also profiled. Qualitative simulated and experimental artefacts are reported in Fig. 6.5(a), artefacts are quantified in Fig. 6.5(b), and simulated and experimental values of time are reported in 6.5(c).

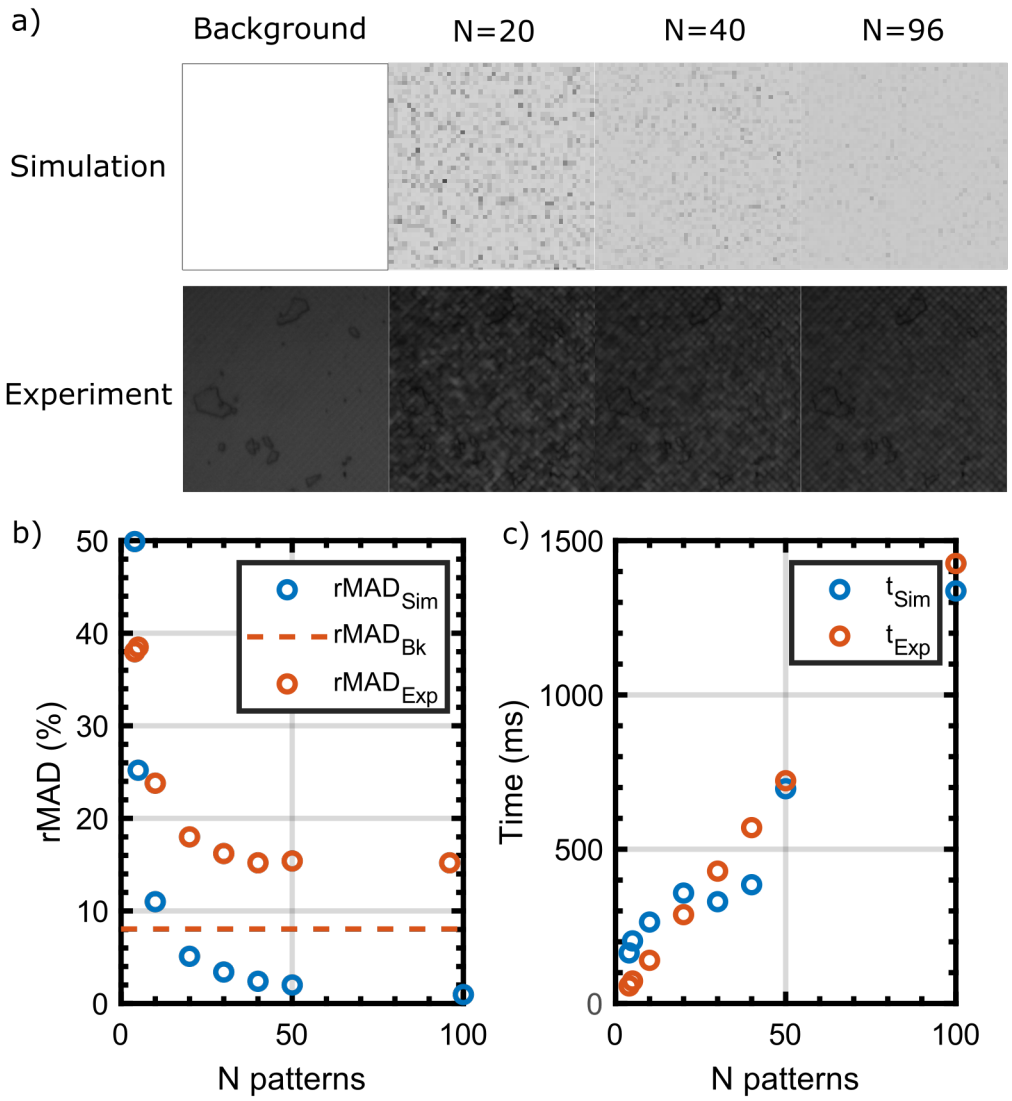


FIGURE 6.5: Comparison of simulated and experimental in-focus artefacts as a function of the number of patterns without a bundle. (a) qualitative comparison based on the number of patterns (N). (b) quantitative comparison using the relative mean of the absolute differences to determine the per cent error as a function of N, Bk - the rMAD of the background mirror, which also shows the pixel edges of the DMD which contribute to the rMAD (c) Time for reconstructions as a function of N. DMD pixel size at object plane 2.5 μm.

Artefacts that result from a DMD are simulated, without a bundle, and then experimentally confirmed (Fig. 6.5). Several conclusions can be drawn. The authors use of 40 patterns in Random Speckle Illumination appears to translate well to the Tetris approach, since there is virtually no artefact improvement adding more patterns.

The results presented in this section were deemed too preliminary to justify tissue validation using animal tissue.

6.4.2 Axial defocus response

Axial response profiles are shown in Fig. 6.6. While the number of patterns does significantly affect the in-focus artefacts and also affects the time taken to acquire images, there is virtually no marginal benefit in terms of improved optical sectioning or defocus artefacts. There is a clear improvement over SIM, with Tetris having a spatial frequency 3 times higher.

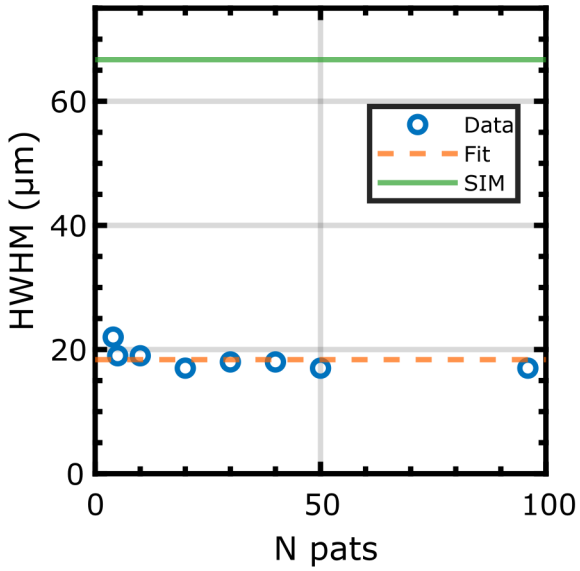


FIGURE 6.6: (a) Half-power point (HWHM) of axial intensity response to defocus. No bundle. Fit: (a) horizontal line at 18.4 μm. SIM axial intensity response with a line pair width 15.0 μm. Tetris pixel size: 5.0 μm.

6.4.3 Bundle effects

A bundle introduces several different effects. First, the size of the DMD pixels at the bundle plane need to be larger than the size of the inter-core spacing of the bundle. Bundle effects become apparent when looking at two limiting cases,

tiny DMD pixels relative to the inter-core spacing and very large. The limiting case of a DMD pixel size much less than the bundle core diameter yields a grey illumination pattern at the distal end. The second limiting case is when the DMD pixels are so large that the entire field of view is either illuminated or not, this offers no added sectioning.

6.4.3.1 Super-pixel size, in-focus intensity and artefacts

At the bundle plane, each DMD pixel (2.5 μm) takes slightly less than the area of 1 fibre bundle core spacing (3.4 μm). To understand the effects this has on imaging a Tetris scheme of 40 patterns with squares of different superpixel sizes is used. The three parameters are then characterised as a function of superpixel size: in-focus intensity, in-focus artefacts, and the HWHM of the axial intensity response (Fig. 6.7). These are then compared to the SIM and WF base cases. The results are quite clear, as the superpixel size increases the in-focus intensity increases. Intensity fluctuation is related to cores averaging over edges. Artefacts approach a local minimum at 29 μm superpixel which does offer an improvement over SIM, and the axial response approaches that of SIM when using 14 μm superpixel.

The inability to obtain improvement the axial response is unexpected since when superpixel dimensions are below 27 μm, the spatial frequency is higher than the 54 μm line pair width of the SIM patterns. There is a clear advantage when using SIM patterns with spatial frequency 3 times lower in the bundle free case (Fig. 6.6), and further improvement is limited with a bundle by sampling effects.

These preliminary results do indicate improved contrast transfer near the ICS

with Tetris. The reconstructed intensity was ~70% of peak intensity for SIM with a LPW 54 μm, and with Tetris of the same spatial frequency (superpixel size 27 μm) it was ~90% peak intensity. The two studies were however done at different times, and there could be a slight error caused by a misalignment in the optical system, or tilt in the fibre bundle against the metallic plate.

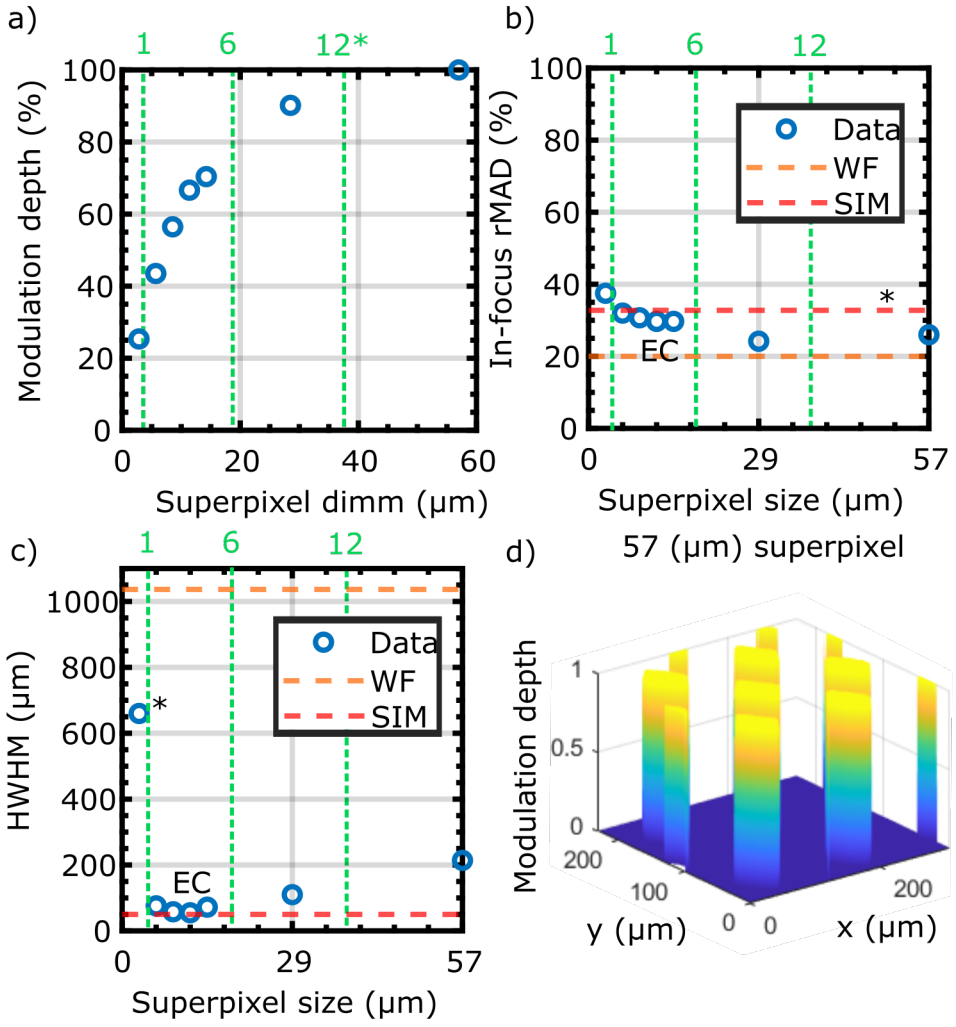


FIGURE 6.7: Tetris parameter optimisation. (a) Modulation depth relative to peak (54 μm), as a function of superpixel width (b) mean of the residuals squared for in-focus reconstructions. (c) axial intensity response to defocus. * indicate ratio of superpixel size to intercore spacing. SIM LPW 54 μm. (d) Four quad modulation 27 μm superpixel, with bundle, raw image normalised by wide-field. EC - equivalent comparison: spatial frequency equal to SIM.

6.5 Four Quadrant

Consider an array of 4 pixels. The Four Quad approach, as the name suggests, uses four images. Choose any arbitrary 4 pixel (or superpixel) square and the location of the illumination across four images will be in each of the four quadrants: [1 0;0 0], [0 1;0 0], [0 0;1 0] and [0, 0;0 1]. The intensity needs to be scaled by an additional factor of 2 since half as many pixels are illuminated. The in-focus image is shown in Fig. 6.8(a). The HWHM was 12 μm; the in-focus MAD was 30 %.

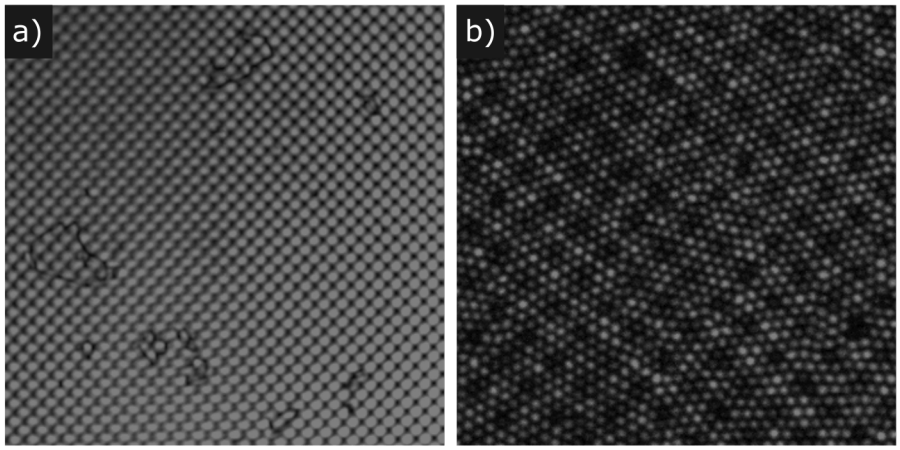


FIGURE 6.8: Reconstructions of an image produced using the Four Quad approach (a) without and (b) with a bundle.

Similar to SIM a feasibility study is conducted. The simulation considers the axial response to defocus of a thin fluorescent mirror moving away from defocus. Relative intensity is determined at the SIM half power-point for Four Quad. Raw patterns are convolved with the kernel which corresponded to the SIM half-power point. The result was a 38% reconstructed intensity for Four Quad. This indicates superior sectioning at the same spatial frequency as SIM. 14% is returned when using a spatial frequency 3 times higher, an early indication of even further performance improvement.

	HWHM (um)	In-focus rMAD (%)
Tetris	18.4	16†
Four Quad	12	30
WF	66	8.5
SIM	18	10

FIGURE 6.9: HWHM of axial intensity response to defocus and in-focus artefacts (rMAD). Tetris $N = 40$ patterns. SIM LPW 14.2 μm. Superpixel size: 9 μm.

Experimentally, axial response to defocus and in-focus artefacts are determined at a single data point of a 9 μm superpixel. The superpixel width was chosen because it is very close to the minimum HWHM of the Tetris approach. Without a bundle, the axial response HWHM for SIM is 18 μm, and Four Quad is 12 μm. With a bundle the values are 22 and 21% respectively, indicating superior sectioning performance. The in-focus artefacts however had an rMAD of 16% with SIM and 29% with Four Quad (Fig. 6.8(b)). The rMAD is due to the halved-core problem which is described below.

6.6 Four method comparison

A comparison table of the two methods, as well as SIM and widefield imaging is presented in Tab. 6.9 when not using a bundle. It is worth noting the Four Quad approach delivered superior sectioning over all other methods, including SIM.

The comparison case with the bundle is presented in Tab. 6.10. Due to unacceptable artefacts and limited sectioning improvement tissue imaging is not currently being explored.

	HWHM (μm)	In-focus rMAD (%)
Tetris	43	18
Four Quad	21	29
SIM_CC	22	16
WF	1040	16

FIGURE 6.10: HWHM of axial intensity response to defocus and in-focus artefacts (rMAD) when using a bundle. Super pixel width 14 μm, SIM LPW 28 μm.

6.6.1 The halved-core problem

When sampling a proximal pattern with a bundle, if a fibre core is halved-by 2 DMD pixels, or superpixels, in each image, it will be sampled at the same intensity. The halving means that during a subtraction based approach, it is automatically subtracted out as out-of-focus light. SIM does not have this problem since the lines shift phase, meaning the edges are not in the same place in all three images.

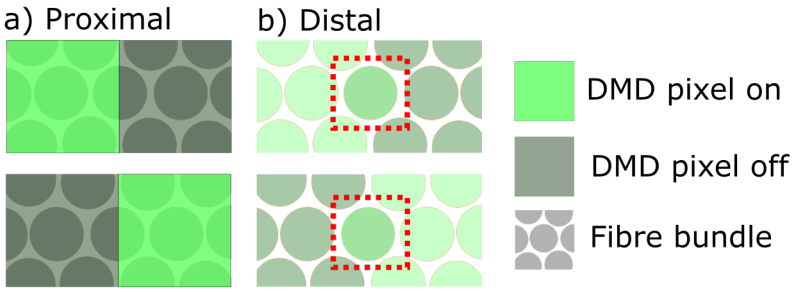


FIGURE 6.11: Halved-core problem. A fibre core on the edge of two pixels is sampled to have the same intensity value among any two frames. Pattern (a) projected onto proximal face, (b) transmitted through bundle.

When using superpixels, it is possible to generate random patterns with pixel grids that are not collocated among images. 40 superpixel patterns are generated with block size 14.4 μm. Each alternating pattern is shifted by 7.2 μm down, and 7.2 μm right. The shift means that no core will be halved in two consecutive runs. The bundle is then placed in contact with a smooth metal plate and acquire images. Each image is then normalised by a low pass filtered version of itself

to correct for intensity variations, the relative mean of the absolute differences is then reported. Qualitatively, there is a vast reduction in artefacts; this can be seen in Fig 6.12. Quantitatively, this leads to a reduction in the relative residuals from 18 % to 16 %.

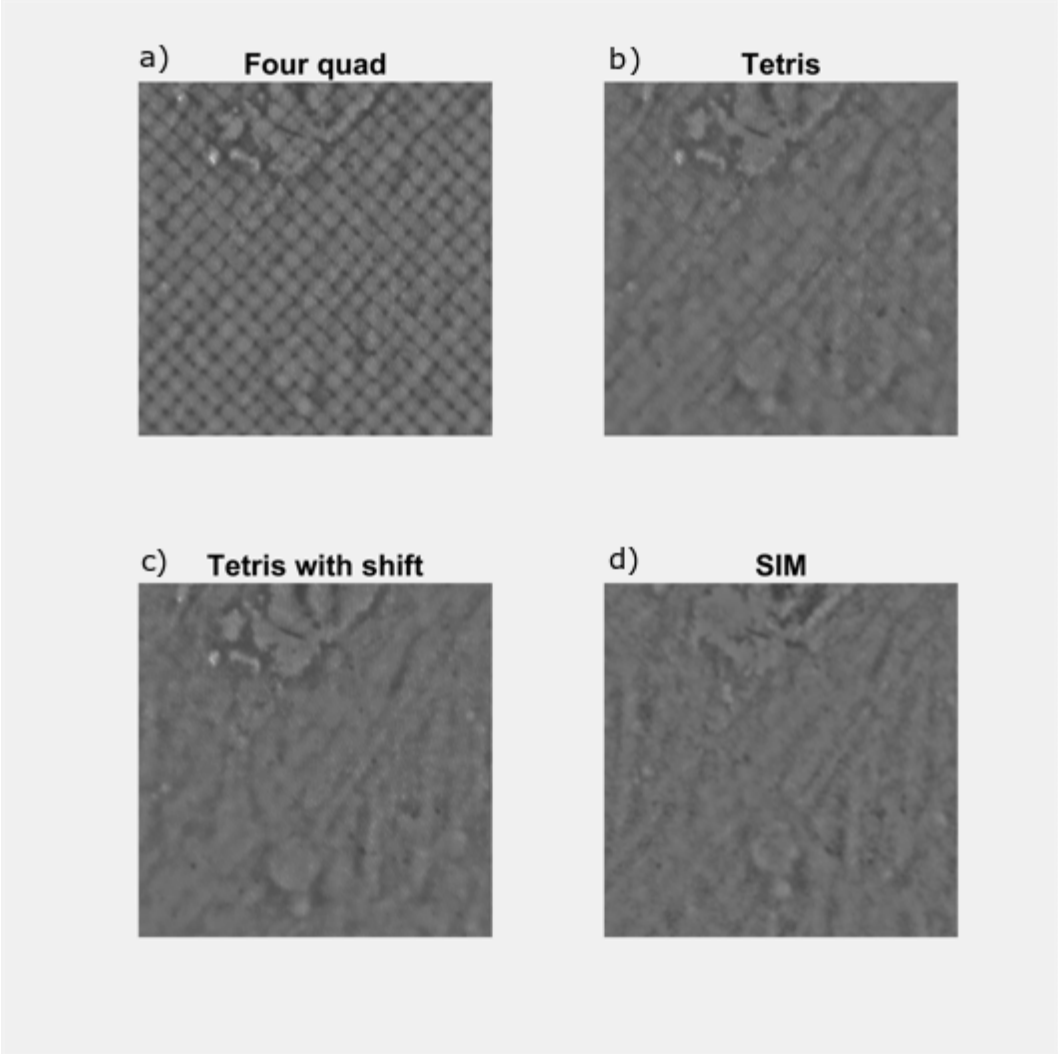


FIGURE 6.12: Bundle in light contact with fluorescently labeled metal plate. Artefact comparison of intensity normalised images. (a) Four Quad, (b) Tetris, (c) Tetris with 1/2 superpixel shift between frames, (d) SIM. Super pixel size 14.4 μm, SIM modulation frequency 29 μm.

This solution, based on the random nature of the detection scheme, only works for Tetris. A solution to the halved-core problem for Four Quad would allow for optical sectioning improvements presented in Tab. 6.10.

6.7 Discussion

When using line pattern illumination, the highest spatial frequency that can pass is limited by the contrast transfer properties of the bundle, as was seen in the SIM section (Chap. 4). If random speckle illumination can have the effect of controlling the random intensity on a core-wise basis, it follows that higher spatial frequencies will pass the transfer function. When using superpixels with sizes near the inter-core spacing, and considering a single pixel's intensity values in two images, with the shifting approach, the expectation value will now always be 50%. Suggesting an ideal frequency transfer nearer to the Nyquist limit than SIM allows. This improved spatial frequency transfer seems to be behaviour that Four Quad is exhibiting, with both superior sectioning with and without a bundle. Problematically, the artefacts are a factor of 2 higher than SIM.

All approaches presented are subtraction based techniques that obtain contrast by taking differences of images. It is necessary that, when imaging in focus features, the illumination intensity through individual cores are different in each frame. The most significant contributor to artefacts when using these approaches is the halved-core problem.

6.8 Conclusion and future work

The preliminary set of results and conclusions here indicate that the bundle translates a greater amount of contrast at higher spatial frequencies using 2D modulation. However, this needs validation with well-controlled conditions (e.g. probe position, room lighting, fluorescent plate). Problematically, as a subtraction based

approach artefacts are introduced caused by the intensity averaging operation, each core performs (the halved-core problem). The artefacts can be further reduced, introducing a pixel shift between Tetris frames. If the halved-core problem could be solved in a similar way for Four Quad, it could lead to improved sectioning with minimal processing time added.

Chapter 7

Enhanced line scan endomicroscopy without artefacts

7.1 Introduction

This chapter presents developments in line scan endomicroscopy (LSE), a fibre bundle based approach based on line scan (LS) confocal microscopy. It presents a new approach to optical sectioning with single frame acquisition. The literature review will introduce relevant existing literature in benchtop (LS) systems, and LSE, systems and subtraction based approaches such as enhanced line scanning (ELS). The theory section of the literature review will demonstrate why ELS delivers improved sectioning performance over LS, by attenuating far-from-focus light.

The subtraction based approach to attenuate far-from-focus light requires two images, one with a detector and illumination slit aligned, and one with the slits offset. This means either two frames are needed, or multiple cameras. When two frames are used this leads to motion artefacts. This chapter describes a novel approach of a single camera configuration to capture two images in a single frame to deliver improved sectioning without multi-frame motion artefacts (ELS+). The expected improvement in axial response when compared to LS is confirmed. An

improved robustness to artefacts is demonstrated by imaging fluorescently labeled lens paper and validated with porcine oesophagus tissue. High speed mosiacs (240 frames/s) are also reported.

7.2 Background

The gold standard approach for endomicroscopy has been point scan confocal, and two commercial systems, the Mauna Kea Cellvizio, and Optiscan Five2 have been developed. Widefield systems such as the HRME have been deployed for clinical use such as treating rural surgical and oral cancers [1, 40]. Confocal point scanning presents advantages over widefield systems in a few ways, the point illumination and detection means that light is only accepted from the focal plane, as a consequence other reflections in the optical system which may have leaked through the dichroic beam splitter or emission filter are reduced [159]. While image quality in these systems exhibits superior optical sectioning performance to conventional. The need for scanning along 2 axis can lead to low frame rates. When scanning distally, the need for a miniaturised scanning mechanism further limits speed [2]. Widefield imaging can be used to obtain images at the full frame rate of a camera and > 100 frames/s with sufficient signal power is realistically achievable. Problematically, widefield acquisition has no added optical sectioning beyond what the bundle itself provides. Line scanning bridges the gap between high speed widefield acquisition, and the out-of-focus rejection provided by point-scanning.

7.2.1 LS benchtop systems

A different confocal principle to point-scanning is the idea of the scanning slit. In 1911 Gulstrand [160] presented his slit lamp. The illumination beam is obliquely incident on the cornea and observed from a direction where only specularly reflected light is imaged [161]. In 1974 Maurice, M., et al. developed the first slit scanning microscope [162]. Ophthalmologists found that when attempting to image an endothelial-cell-layer in the cornea, reflections from the cornea's surface and scattering in its central region and stroma dominated images [163]. The solution proposed was to obliquely illuminate the cells using a thin slit of light in half the objective and collect light in the other half. In 1994 Masters, B., et al, demonstrated a galvonometer scanning based approach which enabled *en face* 2 μm sections of the eye in real-time [161]. As will be seen in the theory section (Sect. 7.2.2), this approach, while fast, is plagued by far-from-focus light degrading images.

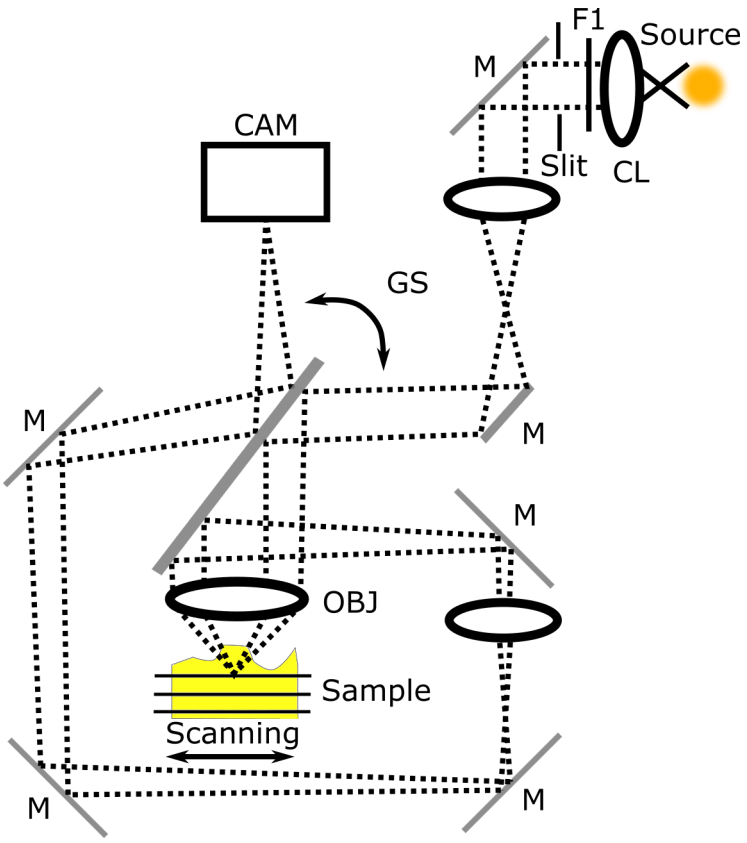


FIGURE 7.1: Line scan benchtop system developed by Masters. A galvanometer scanner is used to control guiding light through the optical system.

7.2.1.1 LS benchtop systems with subtraction

In a pinhole system, the full theoretical resolution improvement can only be realised with an arbitrarily small pinhole. Since axially displaced light is enlarged in a cone over the detector, this would lead to the optimal rejection of axially displaced light since an infinitely small fraction would hit the detector. [164]. In 1998 Gauderon, R. et al., presented an approach to improve the resolution of a confocal pinhole system which can be applied to LS [165]. Since it is not possible to physically create an infinitely small pinhole, and it was found that resolutions in both the lateral and axial directions are improved by subtracting an image formed with light emanating from a large pinhole from that formed with light emanating from a small pinhole. The primary goal of this work was to ensure higher

lateral spatial frequencies passed the OTF. This form of subtractive imaging is like the technique of unsharp masking [165]. Unsharp masking is an approach where a low pass filter of an image is taken, if the difference in the low pass and original images is greater than a specified threshold, the low pass image is subtracted [166]. The approach also has benefits to axial resolution. As will be seen in the theory section (Sect. 7.2.2), as a detector, pinhole or slit, is laterally translated (in the case of the slit, perpendicular to the orientation of the slit), detector intensity, from light returning from in-focus, decays at a much shorter distance than that returning from out-of-focus.

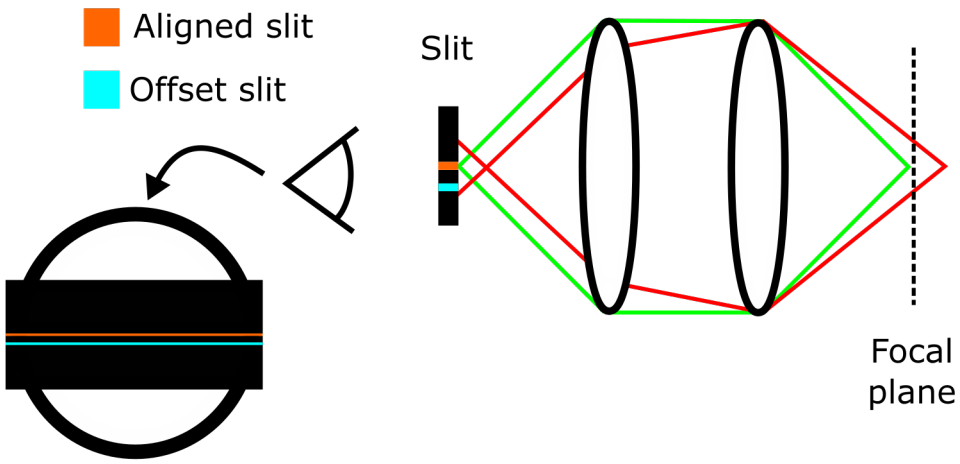


FIGURE 7.2: Principle of ELS. A line emanating from a sample is shown (green), in-focus light returns only to the slit, while axially displaced light (red) is enlarged over not just the slit but a range of lateral slit positions. An offset slit would acquire only light which emanated from an axially displaced location.

This subtraction-based technique was adapted from point-scan to LSM in 2008 by Poher, V., et al. The theory for this is developed in the following section. They reported, under certain conditions, a LS implementation with subtraction was able to achieve slightly better axial resolution than a confocal point scanning system without subtraction. The basic idea is shown in Fig. 7.2. A high percentage

of light returning from an in-focus line returns through the slit, when the slit is translated it would capture very little in-focus light but a comparable amount of out-of-focus light. The optical diagram of a multislit system is shown in Fig. 7.3. Light is guided from a strip LED to the sample via a condenser, excitation filter, dichroic mirror and objective, and returns through the objective, dichroic beam splitter and is imaged onto a camera.

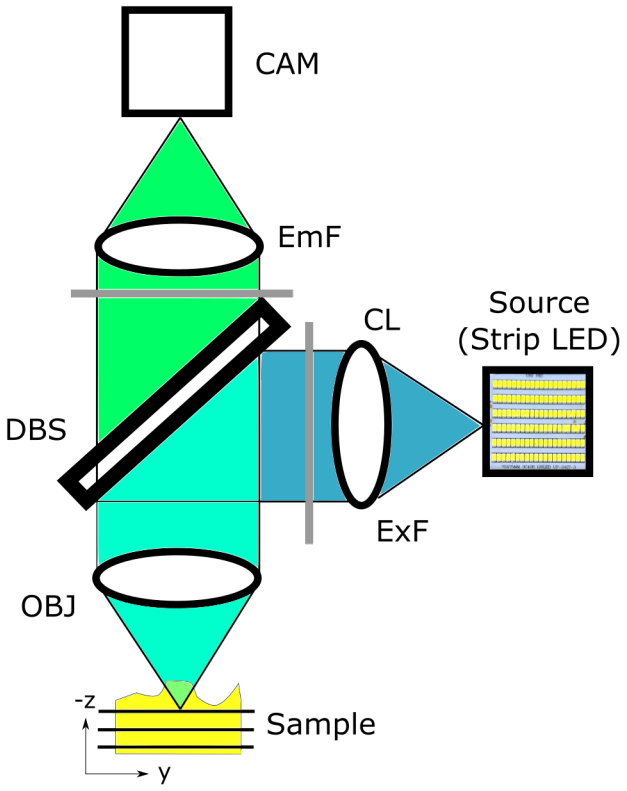


FIGURE 7.3: Line scan subtraction system based on system developed by Poher [82]. Fluorescence system with emission separated by a dichroic beam splitter (DBS), and emission (EmF) and excitation (ExF) filters. Strip LED, conjugate to sample and camera, allows for individual lines to be illuminated. Based on system described in Ref. [82]

7.2.2 Theory of subtraction based LS confocal

The theory in this section is based on that developed by Poher, V, et al. [82]. Effectively, it provides a mathematical basis for the claim that, at points perpendicular

to line in an image plane, the intensity of light emanating a line in the lateral plane far-from-focus is slowly spatially varying, and light emanating from near focus is quickly spatially varying.

The theory uses normalised optical coordinates throughout. The lateral optical coordinate is:

$$v = \frac{2\pi n}{\lambda} x \sin(\alpha) \tag{7.1}$$

The axial is given by:

$$u = \frac{8\pi n}{\lambda} z \sin^2(\alpha/2) \tag{7.2}$$

The normalised optical frequency is given by:

$$s_x = \frac{\lambda}{n \sin(\alpha)} f_x \tag{7.3}$$

where x, z, are lateral coordinates, f_x is frequency, u, is the axial coordinate, v, is the lateral, s_x, is the corresponding normalised coordinate to f_x, λ is the wavelength of light, the NA is the numerical aperture, n is the refractive index, $NA = n \sin(\alpha)$, and f_x frequency perpendicular to the slit. The formula for the response to defocus of a slit scanning microscope is given by:

$$I(u, \beta) = C \int_{-2}^{2} T^2(s_x, 0, u) \text{sinc}\left(\frac{p s_x}{2\pi}\right) \text{sinc}\left(\frac{q s_x}{2\pi}\right) \times \exp -i s_x \beta d s_x \tag{7.4}$$

where, T is the incoherent transfer function, p and q are, respectively, the slit widths of the illumination and detection slits. β is the distance between the illumination and detection slits, all in normalised lateral coordinates. The incoherent optical transfer function (Fig. 7.4) is given by:

$$T(s,u) = \begin{cases} g(s)2\left[2\frac{J_1[us(1-s/2)}{us(1-s/2)}\right] & \text{if} 0 < s < 2 \\ 0 & \text{if} 2 \leq s \end{cases}$$

where $s = \sqrt{s_x^2 + s_y^2}$. The assumption is that the OTF is identical for illumination and detection paths (no Stokes' shift in the fluorescence process). Stokseth's approximation is used to estimate the variation in the OTF with defocus:

$$g(s) = 1 - 0.69s + 0.0076s^2 + 0.0437s^3 \tag{7.5}$$

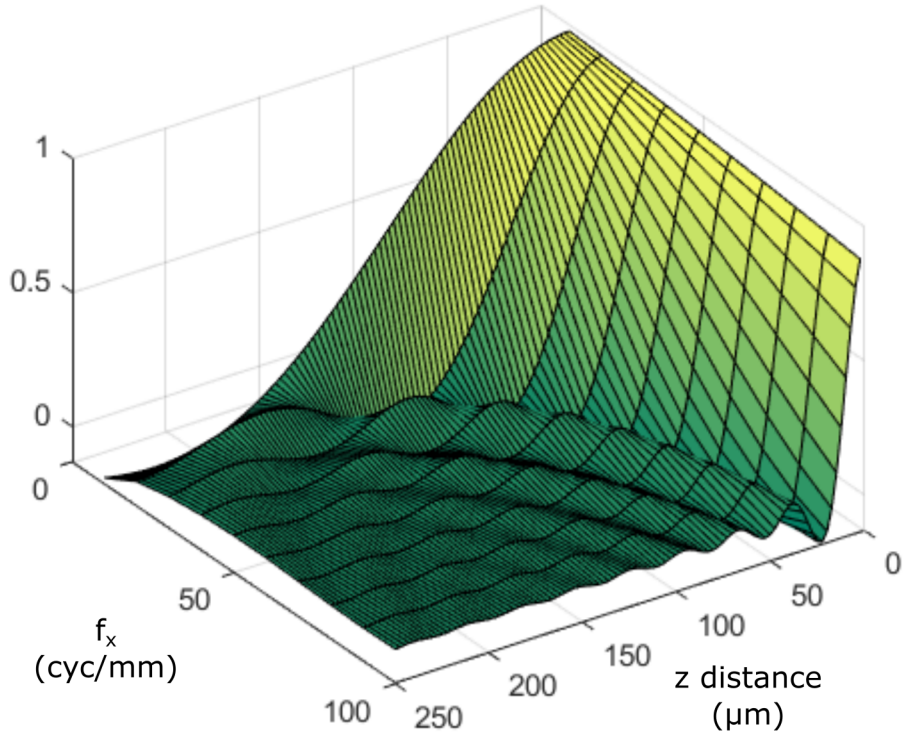


FIGURE 7.4: OTF curves as a function of z distance and spatial frequency.

The in-focus intensity drops off quickly with slit displacement, and out-of-focus light drops off slowly. This can be seen by substituting realistic experimental parameters into Eqn. 7.4 and numerically integrating. Intensity values, relative

to the aligned-slit case, are determined with a 3 μm detection and illumination slit widths, a defocus of $z = 0.5$ μm, an NA of 0.39 (an estimate to that of the fibre bundle used in experiments), $\lambda = 488$ nm, p and q of ≈ 5 μm. For the in-focus case, at seperations of 3 μm, 6 μm, and with a 12 μm separation, the intensities are respectively, 68%, 37%, and $< 1\%$. The defocus case drops off much less sharply with slit separation. This can be seen with defocus of $z = 100$ μm. Relative to the aligned slit intensity, with 3 μm, 6 μm, and with a 12 μm separation, the intensities are 99%, 97%, and 91%.

An image acquired with an aligned slit has contributions which are in-focus and out-of-focus:

$$I_{\beta=0} = I^{in}_{\beta=0} + I^{out}_{\beta=0} \tag{7.6}$$

since out-of-focus contributions are slowly varying we can write:

$$I^{out}_{\beta=0} \approx I^{out}_{\beta=-\epsilon} \approx I^{out}_{\beta=+\epsilon} \tag{7.7}$$

ϵ is a small lateral offset, chosen to be in a region where the in-focus signal, relative to $\beta = 0$, is very small. We can now subtract out the out-of-focus contributions using:

$$I_{\beta=0} = I^{in}_{\beta=0} - \left[\frac{I^{out}_{\beta=-\epsilon} + I^{out}_{\beta=+\epsilon}}{2}\right] \tag{7.8}$$

By choosing a symmetric beta, after substitution of Eqn. 7.4 into Eqn. 7.8, we can write the defocus response of our ELS system as:

$$I_{ELS} = I(u) - \frac{I(u,\beta) - I(u,-\beta)}{2} \tag{7.9}$$

This implies:

$$I_{ELS} = C_s \int_{-\infty}^{\infty} T^2(s_x, 0, u)\text{sinc}\left(\frac{ps_x}{2\pi}\right)\text{sinc}\left(\frac{qs_x}{2\pi}\right) \times [1 - \cos(s\beta)]ds_x \tag{7.10}$$

where C_s is a constant. This can now be compared to the point-scan confocal case, and the slit scan, without subtraction. The formula for point-scan is:

$$I_{PS} = C_l \int_{0}^{\infty} T^2(s, u)\text{jinc}\left(\frac{ps_x}{2}\right)\text{jinc}\left(\frac{qs_x}{2}\right) sds_x \tag{7.11}$$

where C_l is a constant and $jinc(x) = J_1(x)/x$. p and q are the diameters of the illumination pinholes in the point-scan case. The formula for slit scan, is simply Eqn. 7.4 with $\beta = 0$:

$$I(u, \beta) = C_l \int_{-2}^{2} T^2(s_x, 0, u)\text{sinc}\left(\frac{ps_x}{2\pi}\right)\text{sinc}\left(\frac{qs_x}{2\pi}\right) ds_x \tag{7.12}$$

C_l is a constant. By substitution of realistic experimental values, and numerical integration, axial response curves were generated for the three methods. Substituting in each of the formulas (Eqns. 7.4, 7.12, 7.11) experimentally practical values of $p = 3$ μm, $q = 3$ μm, $\beta = 12$ μm. NA = 0.39, $\lambda = 488$ nm yielded the curves shown in Fig. 7.5.

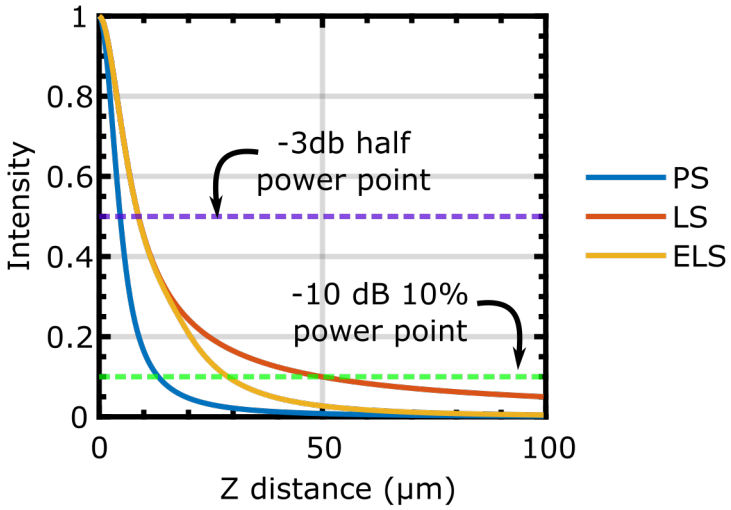


FIGURE 7.5: Theoretical axial response curves using experimentally similar values. $p = 3$ μm, $q = 3$ μm, $\beta = 12$ μm. NA = 0.39, $\lambda = 488$ nm.

The half-power point as well as the 10% power point are identified as horizontal lines. With the realistic experimental values used there is a predicted reduction of the 10% power point using a subtraction based approach. Experimentally varying the offset as well as slit width are studied in Sect. 7.3.1.1. Far-from-focus the intensity of a LS system decays as $\frac{1}{r}$, wheras, point-scan confocal falls off as $\frac{1}{r^2}$. When summing images acquired at different depths, the intensity will fail to converge [159]. This can be corrected as part of the multi-frame subtraction based approach.

7.2.3 LSE with a fibre bundle

An adaptation of a benchtop LS approach [167] was applied to endomicroscopy, without subtraction, by Sabharwal, A., et al. in 1999 [168]. The optical system is shown in Fig 7.6. The principal is an Argon Ion laser is focused to a line by a cylindrical lens. That line is then relayed via a scanning mirror to a conjugate focal plane where the bundle is placed. The returning light is then guided down

a separate path, where there is a detection slit aperture on a conjugate focal plane which filters out-of-focus light. That light is then guided via mirrors and the scan mirror onto the surface of a camera. Based on geometry the scanning mirror is used as both the detection mask as well as illumination slit.

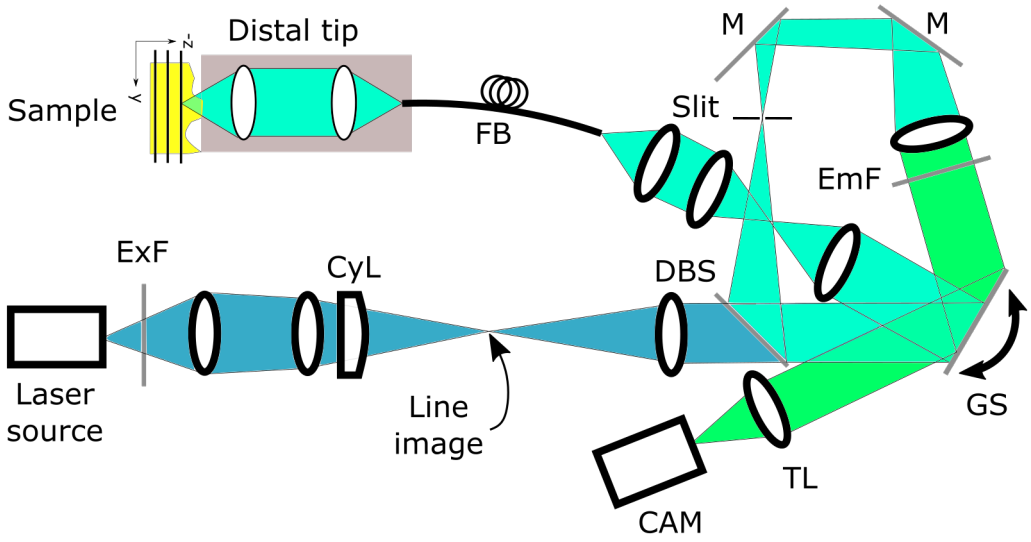


FIGURE 7.6: Fluorescence line scan endomicroscope. Laser source, focused to a line by cylindrical lens (CyL), line image conjugate to fibre bundle (FB) and slit. Fluorescence system with emission separated by a dichroic beam splitter (DBS), and emission (EmF) and excitation (ExF) filters. Galvonometer scanner (GS) used to scan line. Based on system described in Ref. [168].

This approach requires fine alignment of several components. The system also needs to be optically isolated from the room light to avoid stray reflections. The FWHM of lateral and axial resolutions was 3 and 25 μm respectively. The system used distal lens elements. The field of view was 430 μm. The rigid diameter is 7.9 mm, and the rigid length is 60 mm. The maximum frame rate was 4 frames/s.

In 2015 Hughes, et al. [78] reported an improved system which used a linear CCD positioned where the detection slit was in Fig 7.7. This improved speed and the authors incorporated mosaicking. The system is similar to Sabharwal's design. The lateral and axial resolutions were 2.2 μm and 10 μm (Estimated from

axial response profile). The speed was 120 frames/s. The field of view was a diameter of 250 μat the samples focal plane, explaining the trade off in field-of-view (FOV) and resolution. The study was done with a Mauna Kea probe, since resolution is a function of magnification, the resolutions obtained with a lensless probe, such as the ones determined this chapter, are not comparable.

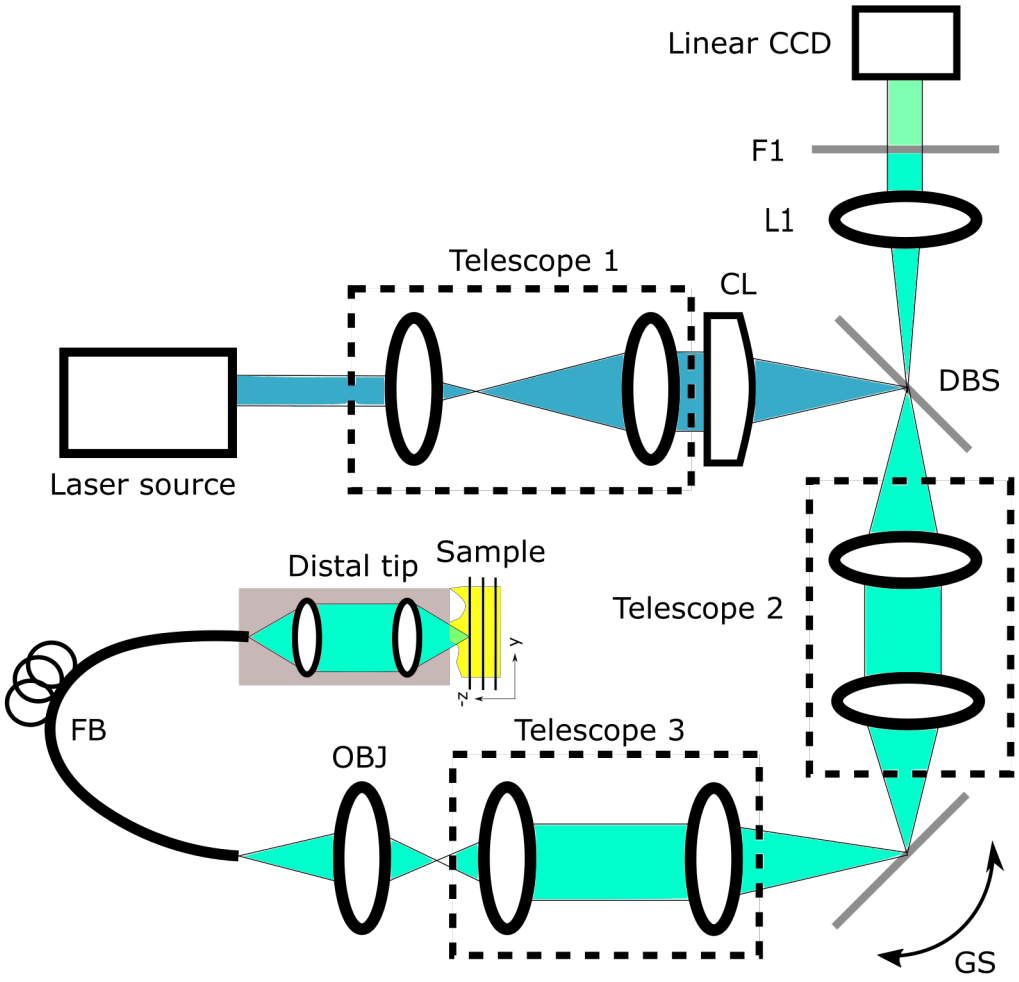


FIGURE 7.7: High speed line scan endomicroscopy system developed by Hughes. Laser source, focused to a line by cylindrical lens (CyL), line image conjugate to fibre bundle (FB) and linear CCD. Fluorescence system with emission separated by a dichroic beam splitter (DBS), and emission (EmF) and excitation (ExF) filters. Galvonometer scanner (GS) used to scan line across bundle face. Based on system described in Ref. [78]

In 2016 the idea was expanded [79] and the high speed linear camera was replaced with a rolling shutter camera of the same frame rate. The rolling shutter

camera reads out rows of pixels sequentially. When each row, or sequence of rows is read out, the rows above and below are inactive, and light is not collected there. This resulted in physical blocking of offset light in the same way as the physical slit. The benefit is the slit width is electronically variable, and controlled by the exposure. This approach will be used in the work in this chapter, and is discussed in the method (Sect. 7.3.1.1). This camera also ran at the full frame rate.

The same problem occurs in LSE as LSM, light far-from-focus degrades images this is due to the decay of intensity as $\frac{1}{z}$ as opposed to the $\frac{1}{z^2}$ with point-scan confocal. Hughes, et al., adapted Poher's, et al. subtraction based approach to LSE. In LSM, there is no sample movement and having multiple-frames does not degrade images. While in LSE, the probe is often moving and as a multiframe approach it is subject to motion artefacts.

7.2.4 LSE with a fibre bundle and a DMD

A DMD for point-wise confocal endomicroscopy has advantages in that each mirror of the DMD array can be selected electronically, this allows for patterns to be customized, the main Literature Review (Chap. 2), shows an example of this where the authors used multiple DMD pixels to select individual cores on a bundle. That approach used the cores of the bundle as illumination pinholes, and in software mapped the core locations to the camera pixels. This mapping allowed for a software mask of the surrounding cores. The illumination and masking protocols are shown in Fig. 7.8. With 30,000 cores speed is a major limiting factor, so the authors used multiple illumination and detection cores. Parallel detection of light has also been reported for non-endomicroscopy applications [77].

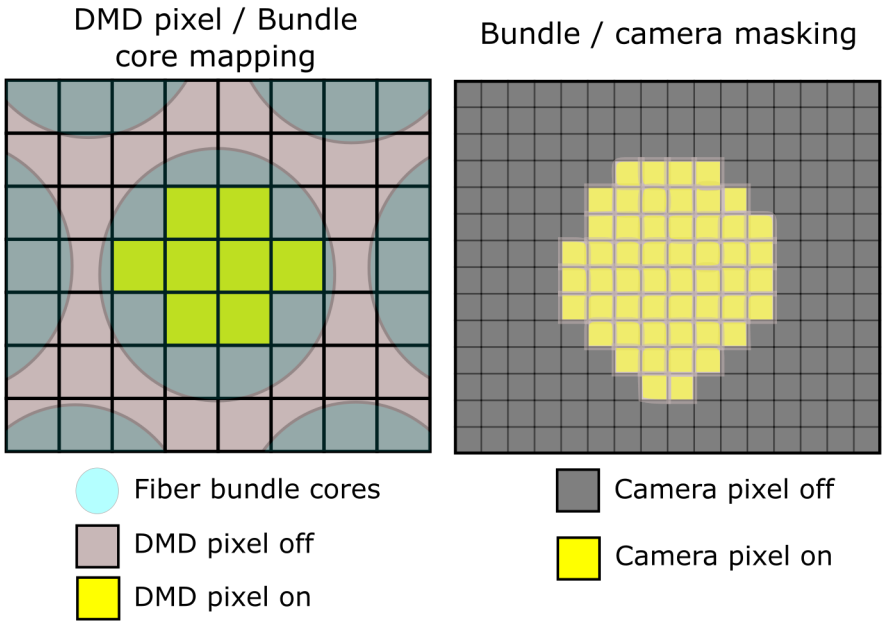


FIGURE 7.8: Single core of a fibre bundle illuminated by multiple DMD pixels. During calibration, an image reflected off a mirror is acquired and the camera pixels are used to generate a mask in software. Reproduction of principle presented in Ref. [76].

There are limitation to the core wise approach which needed to be overcome for real time imaging, including parallel collection of light from cores, which requires multiple frames, introducing motion artefacts. Another way to use a DMD is as an electronically adjustable illumination slit.

Tang et al. presented in Nov 2017, a configuration of a fluorescence DMD and rolling shutter camera system [80]. The electronically controllable mirrors of the DMD were used as the illumination line, while the rolling shutter of the camera was used as the detection line. The schematic is shown in fig 7.9. Light is directed from the blue LED of the DMD, reflected off the mirror array, collimated by a 120 mm lens, reflected off of a dichroic beam splitter mirror and onto the tip of a fibre optic bundle. The pixel array of the DMD is aligned with a camera's rolling shutter array. The device used a lens free bundle in contact mode for imaging.

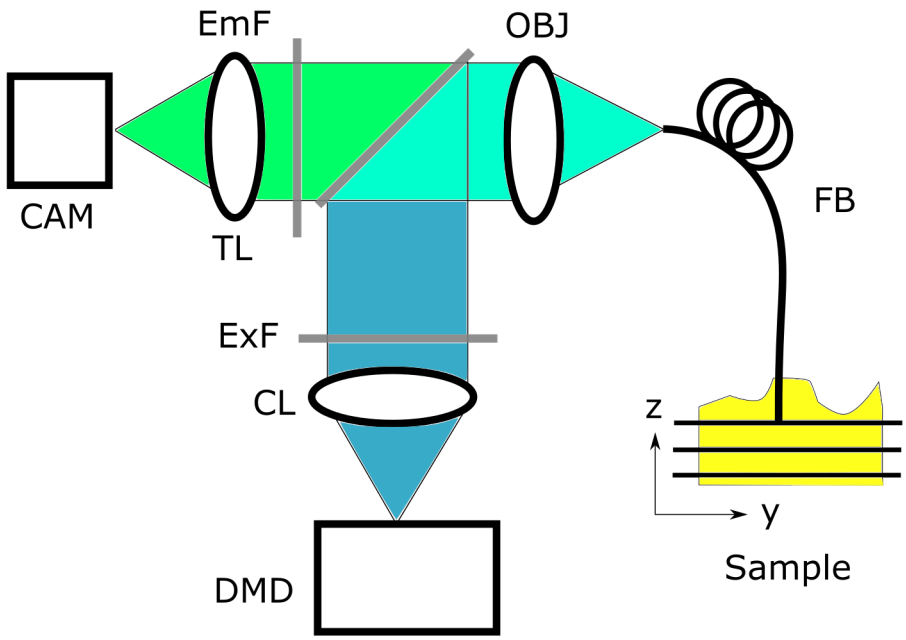


FIGURE 7.9: LSE with a digital micromirror device. A line image is generated by imaging a DMD, illuminated by the built-in LED onto a fibre bundle (FB). The returning fluorescence is then imaged onto a rolling shutter camera. Fluorescence system with emission separated by a dichroic beam splitter (DBS), and emission (EmF) and excitation (ExF) filters. Based on system described in Ref. [83].

Since the rolling shutter camera reads out lines at a set rate and the entirety of the camera's CCD must be swept at 120 Hz, this means the rolling line of a DMD must also "sweep" at the same rate. The DMD the authors used could only change patterns at a much slower rate than this, making a single DMD line, mapped exactly to a single camera line, unfeasible. Consequently, the authors developed an approach where larger illumination lines were used. During this spatial gating process the detection line swept across the large illumination gates, the timing was such that once the detection had finished its sweep, the next series of illumination gates was opened and previous closed.

This has a few complications that are illustrated in Fig 7.10. The first complication is from incomplete collection, and the second from inconsistent collection of out-of-focus. The incomplete collection resulted from where the detection

line started its travel. In that case the detection lines and illumination lines only slightly overlapped at the top and bottom of the travel. The inconsistent collection of light which was out-of-focus however, is more subtle. Consider the detection line sweeping across a single illumination line gate with the detection lines travel exactly in the center of the illumination line. At this point it is capturing both the in-focus light and out-of-focus light from both sides of the detection line. However, at the top and bottom of the travel it is capturing the in-focus light and 50% of the multiply scattered light.

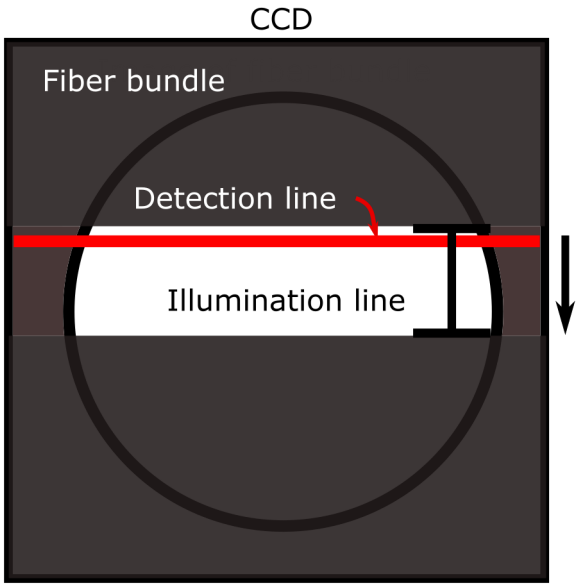


FIGURE 7.10: Line scan sequence for a single illumination gate over the proximal face of fibre bundle. Detection line reads out at 120 Hz, pattern is changed once the detection line reaches the end of the gate. Once end of travel is reached, the illumination gate shifts to the top of detection line. Based on procedure described in Ref. [81]

The authors resolved the incomplete collection problem by collecting two images, with the illumination gates offset by 50%, the two images were then averaged to reduce the line artefacts. This appears to reduce the multiple scattering artefacts as well as they are averaged with artefact free regions of the sample. The

spatiotemporal diagram (Fig. 7.11) shows the collection of the two images. Consistent with theory of LS confocal, optical sectioning performance exhibited better performance with more narrow illumination slits. The removal of artefacts was solved by adding additional frames, making it susceptible to motion artefacts.

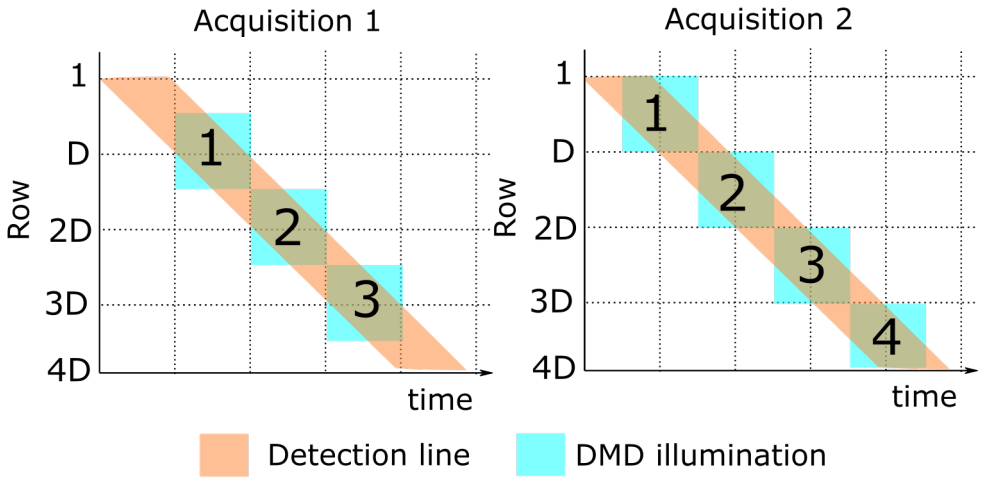


FIGURE 7.11: Spatio-temporal distribution of DMD illumination and CMOS collection. Numbers represent sequence of illumination, DMD row D – represents the size of a single illumination gate. Based on system described in Ref. [81]

In 2019 Tang et al. published two further papers, the first was designed to overcome to limited pattern storage on the DMD[81]. The second modified the spatiotemporal illumination approach to remove artefacts[83].

In the first paper, the optical system is identical to Fig 7.9. The key difference is the author using parallel patterns to overcome the speed limitation of the DMD. The parallel patterns mean that more narrow illumination lines near focus can be used, since more lines can fit on a single DMD array. They do of course come with a drawback, the multiple regions in sample is simultaneously illuminated, even parts where imaging is not taking place. This can lead to photobleaching. Also, the tunability of the line widths is determined by the scattering properties of the sample. A transparent sample with features far from focus likely would

not be amenable to this approach since the parallel lines away from the detection line would blur through the sample and return as far from focus light. This is particularly problematic in a LS configuration where far from focus light is well known to degrade image quality [159], and can be seen as a long tail in the axial response profile 7.5.

The authors further report a complete removal of artefacts which came from partially overlapping detection and illumination lines. They did this by modifying the line illumination to begin its travel completely overlapping the line illumination. When the next illumination line was displayed, it did not start at the bottom of the detection line but at the top. Fig 7.12 illustrates this principle. While this may have addressed artefacts from overlap, it does not address artefacts that result from multiple scattering at the top and bottom of the travel. The original approach left line artefacts in the images, and the authors needed to collect two images and average them to remove them. The major benefit of the modified approach is the elimination of these artifacts, hence making a multiframe approach a single frame approach. The cost is the illumination line width doubles in size, meaning more out of focus light is collected.

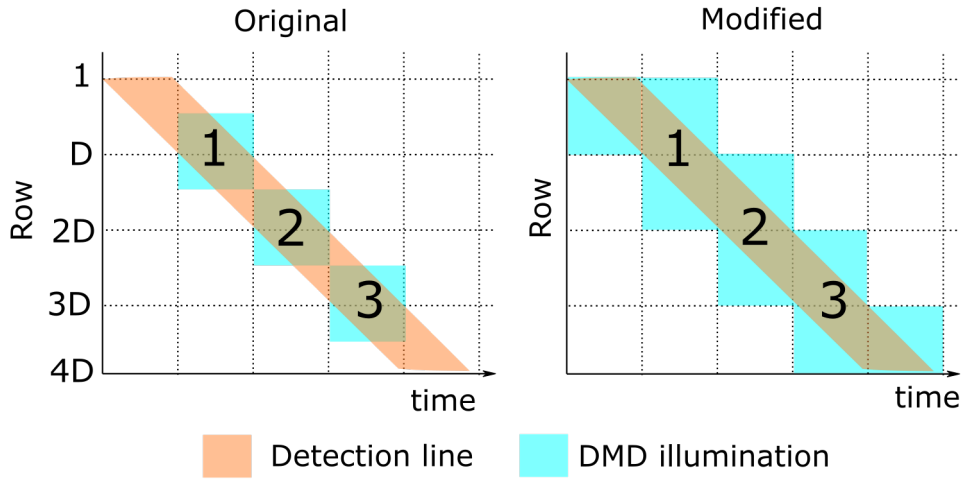


FIGURE 7.12: Modified spatiotemporal diagram, as reported by Tang, to reduce artefacts.

The system presented in the second paper was designed to reduce the effect of out of focus light through a subtraction-based approach similar to Poher, Hughes [79, 82]. It is shown in Fig. 7.13. The authors take the following two frame approach: acquire an image with a narrow illumination slit and a second with a large detection slit. For both frames the illumination slit is the same size. The scaled image with the large detection aperture is subtracted from that with the narrow detection slit. This enables the removal of out of focus contributions.

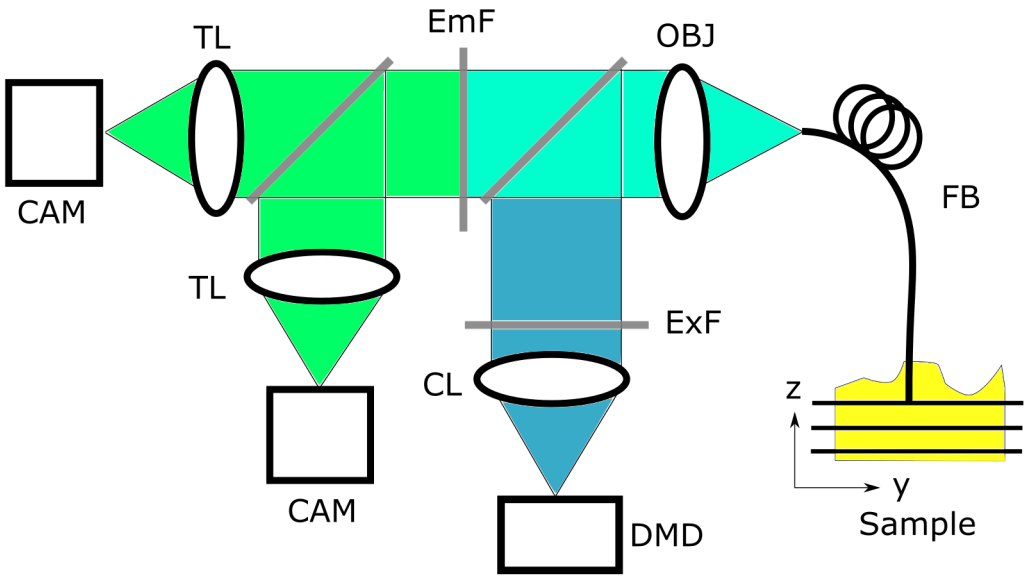


FIGURE 7.13: Line scan with a DMD and two cameras for reduced motion artefacts.

The electronic shutter acts as a detection slit with an adjustable slit-width. The slit progresses at a constant rate however, the detection slit width size is set by the exposure. In practice the exposures of Cam 1 and Cam 2 are set to provide the user the desired slit width.

7.2.5 Summary of prior work

All systems have in common that images are formed by scanning an illumination line over the proximal end of a fibre bundle then the fluorescence emission returning from the sample is imaged onto a detection slit, conjugate with the illumination line, then onto a camera [159]. The first demonstration was with a physical slit and 2D camera by Sabharwal[168]. A faster configuration was then reported using a linear CCD[78], and then using the rolling shutter of a CMOS as an electronically-variable slit with a width that can be changed on-the-fly while achieving a frame rate of 120 frames/s [79]. The problem with LS is light far-from-focus degrades the image. To mitigate this, a two-frame subtraction based approach, previously shown for benchtop systems [82, 169], was adapted to endomicroscopy [79]. The first frame is acquired with the rolling shutter aligned with the galvanometer scanners, the second with an offset between line and detector slit. The offset slit image is then subtracted from the aligned to generate an enhanced line scan (ELS) image. Since two frames are needed this leads to a halving of the frame rate and, as with any multi-frame approach, motion artifacts when the probe shifts between acquisitions. A two-camera approach has been proposed which avoids interframe artefacts[83]. To generate the illumination line, a digital micromirror device (DMD), a programmable array of mirrors is used. The rolling shutters of each camera are used as the detection slits. This approach is also feasible with a laser line and galvanometer scanners. The limitation of this approach is the DMD can only store a limited number of patterns onboard, and to overcome this, parallel lines are used. This has a few drawbacks: first, the width of the illumination lines is constrained by the number of lines which can

fit onto the DMD array, second, it limits the size of the enlarged detection slit, as well as the illumination of regions where light is not being collected which may contribute to scattering and photo-bleaching. A two-camera approach with galvanometer scanners is feasible but would require the second camera, as well as fine alignment.

We now report a single camera configuration of the ELS approach for a galvanometer based system which restores the halved frame rate and eliminates motion artifacts caused by the subsequent frames. We accomplish this projecting images of a bundle onto different parts of a camera and using the rolling shutter of a camera as both the aligned, and offset detection slit (ELS+). This has two benefits, first, it reduces artifacts, especially important in a system with lower frame rates. Second, it allows the full frame rate of the camera to be used. We report system speeds of up to 240 frames/s with lateral resolutions of 8.8 μm, axial resolutions of 10 μm at the half-power point, and 50 μm at the 10% power point, and real-time mosaicking. This is an improvement in speed over previously reported subtraction based single-camera approach (60 frames/s) by a factor of four and the two-camera DMD approach (15 frames/s) by a factor of 16.

7.3 Method

The rolling shutter of a camera will be used as a detection slit. The novel single frame approach will involve using a beam splitter and tilted mirror configuration, on the detection arm, to guide light to the left and right halves of a camera. The tilted mirror will also control the vertical position on the array, this will determine

the slit offset (β in Eqn. 7.4).

7.3.1 System description

A diagram of the new system is shown in Fig. 7.14(a). The spot from a laser diode (ThorLabs L488P60) is collimated through a 4x finite conjugate objective then focused onto a galvanometer scanner (ThorLabs GVS001), then onto the back aperture of a 10x infinity-corrected objective (Nikon PLAN N RMS10X) via a cylindrical lens (ThorLabs LJ169SRM-A), 50 mm achromatic lens (Thorlabs AC254-050-A-ML), a dichroic beam splitter (ThorLabs DMLP490), then relayed onto the face of a 30,000 core fibre optic bundle (Fujikura FIGH-30-650S). The fibre probe employs no distal optics and the probe is operated in contact mode only. The bundle's diameter is 650 μm and the average spacing between the bundle cores is 3.2 μm.

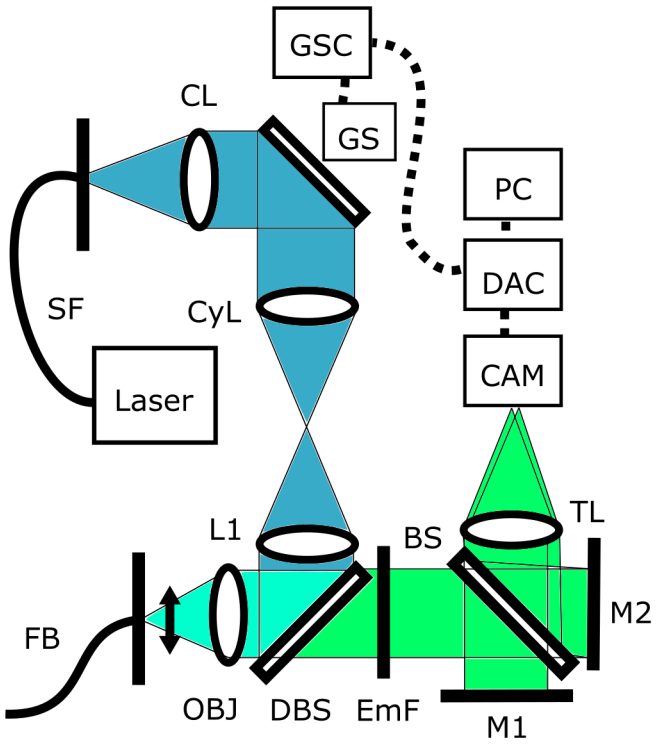


FIGURE 7.14: Single frame ELS schematic. CL - 4x finite conjugate objective, CyL - 50 mm cylindrical lens (power axis only shown), L1 - 50 mm lens, OBJ - 10x infinity corrected objective, TL - 40 mm tube lens, GS - galvanometer, GSC - galvanometer controller, SF - single fibre, FB - fibre bundle (scanning direction across bundle shown), DBS - dichroic beam splitter, BS - beam splitter, M1 - stationary mirror, M2 - adjustable tilt mirror.

The returning fluorescence emission is guided from the bundle through the objective and dichroic beam splitter and is split into two paths by a beamsplitter onto a fixed mirror and a tilted mirror (Fig. 7.15). The mirror configuration is designed to guide identical images of the bundle onto the left and right regions of a camera (FLIR FL3-U3-13S2M-CS) via a 75 mm tube lens (Thorlabs AC254-075-A-ML). One axis of the tilted mirror guides light to the right and left halves of the camera, the second axis controls the offset. Two images of the bundle are shown on the camera plane in Fig. 7.15. The offset of the second bundle image is controllable with the tilted mirror, allowing the rolling shutter to act as both the aligned and offset detection slit. The galvanometer scanners are synchronized with the readout of the rolling shutter using the calibration procedure described below. The magnification between the sample and the camera was 3.8x, the minimum slit width, determined by the dimensions of a pixel was 3.63 μm at the camera plane, and 0.96 μm at the bundle plane.

The offset can not be easily changed while imaging since the new position of the second bundle image on the camera needs to be located. The widths of the detection slits must also be the same. This means subtracting an enlarged slit is not possible with this approach in a single frame. In order to minimise differences in optical paths of the two bundle images, and minimise the aberrations that occur when light goes through different parts of a lens, the camera is on a translating mount. This also helps with high speed (240 frames/s) imaging, since

the bundle images need to be located on a small region of the camera. Ordinarily, it would make sense to fill as much of the camera as possible with a bundle image, however, with this approach clipping effects can be introduced when the offset bundle is clipped at the edge of the camera, this limits the maximum offset possible. This didn't overwhelmingly contribute practically. With a 30,000 core bundle, oversampling was available when filling $\sim 75\%$ of the cameras height with the bundle images, allowing for a shift of up to 200 μm ($\beta = 1,000$). At 100 μm offset ($\beta = 500$) there is no longer any improvement over LS meaning a 200 μm maximum offset is more than sufficient.

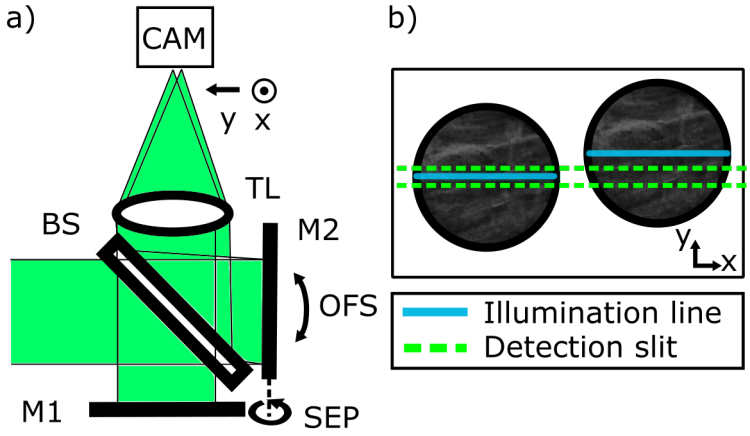


FIGURE 7.15: Tilted mirror configuration. Light from the detection arm is split toward two mirrors and guided onto the left and right halves of the camera. M1 - stationary mirror, M2 - tilted mirror, OFS - offset tilt control, SEP - separation control.

Imaging can be performed using either high sampling or low sampling modes. High sampling can operate at speeds up to 120 frames/s and uses a large portion of the camera area to sample the bundle at 6.8 camera pixels per fibre core with 5.5 μm resolution. Low sampling takes advantage of the high frame rates made possible by using a smaller area of the camera and operates at 240 frames/s and

can sample the bundle at 2.4 camera pixels per fibre core with lateral resolution 8.8 μm.

7.3.1.1 Detection slit parameters

The rolling shutter camera uses sequential line read out, which progresses at a rate of 7.63 μs per line. The exposure determines how many lines are active simultaneously. The relationship between active lines and exposure is

$$\mathrm{N} = \frac{\mathrm{exposure}}{7.63\,\mu\mathrm{s}} \tag{7.13}$$

where N is the number of active lines.

7.3.2 Calibration

The system is calibrated using a custom-built LabVIEW VI (National Instruments, Austin, TX) and consists of three steps: setting the offset to the desired value, mapping the galvanometer to the rolling shutter, and locating the bundle images on the CMOS. This calibration only needs to be performed once for each offset.

7.3.2.1 Offset

The first step is to set the offset to balance optical sectioning strength and signal strength. The offset calibration is shown in Fig. 7.16(a-c). Since the offset will be perpendicular to the laser line, the laser line reflection is used as a reference point. The difference in vertical location between the two line reflections gives the offset. First, the galvanometer voltage is set to 0 V, at this voltage the laser

line appears approximately in the centre of the bundle. Next, an intensity profile is plotted along two vertical lines, one for the aligned bundle, and one for the offset. A threshold filter is applied to remove the background. Each profile is then fit with a Gaussian, where the mean is the central vertical pixel position in high precision, and the difference in the mean of each Gaussian is reported as the offset between the slits. This is updated in real-time to allow instant feedback.

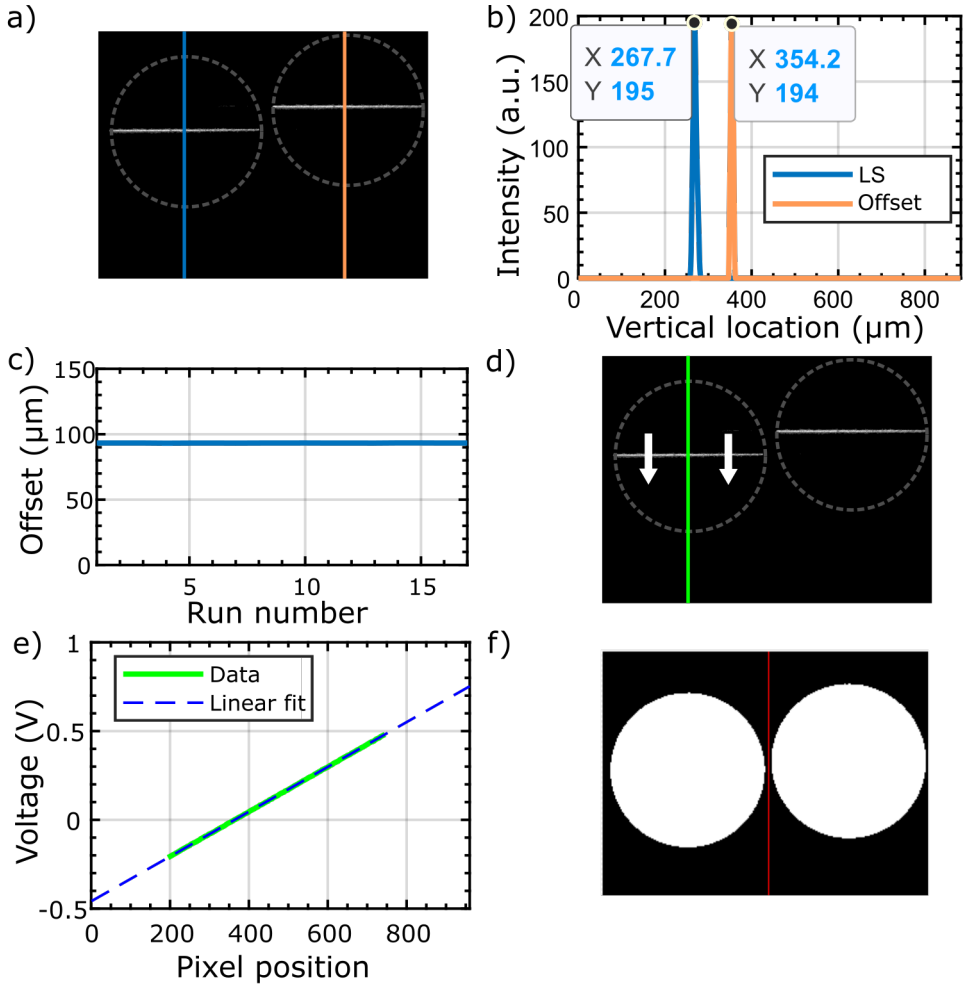


FIGURE 7.16: System calibration. (a-c) Calibration of line offset. (a) High power laser light reflected off the proximal face of a fibre bundle, aligned image (left), and offset image (right). (b) The intensity profile was taken with vertical lines at 1/3 and 2/3 the width of the camera. Gaussian fit to each intensity profile and (c) difference of means reported in real-time. (d-f) Mapping of galvanometer voltage to pixels on CMOS. The line is swept across bundle face (d) and at each step, an intensity profile generated and (e) voltage as a function of pixel position is plotted. (f) threshold mask of distally illuminated bundle image.

A line is fit with the slope determining a mapping between voltage and pixels, the intercept is the starting voltage. Representative mapping: 0.00126 (V/pixel), offset: -0.459 (V), $R^2 = 0.99$. (f) Bundle locator routine. The bundle is distally illuminated with a cell phone LED, and threshold is applied. The centers of the two bundle images are returned as the rows and columns of the bundle with the greatest number of pixels.

7.3.2.2 Laser line to camera mapping

The second step (Fig. 7.16)(d-e) is a fully automated mapping of the laser line to the CMOS in pixels. This is performed by sweeping through at least 400 voltage values, and capturing an image at each voltage, if a line is present in the left half of the camera (aligned slit), then an intensity is taken profile across the line to locate the index of the highest intensity pixel and plotting voltage vs pixel number. A line is then fit with the slope giving the linear coefficient between the galvanometer voltage and camera pixels in Volts/pixel, and the intercept giving the line offset determined in Volts.

7.3.2.3 Locating bundle

The third step (Fig. 7.16(f)) is locating the bundle images on the CMOS. First, the bundle is pointed towards a light and an image is captured. The images are then separated into left and right halves and a threshold mask is then generated. The threshold mask is used to identify the location of the circles. The threshold value should be chosen to include only core and cladding pixels and minimise noise. The next step is to identify the midpoint in each circle. Rows and columns are

handled separately. To identify the central row, we first sum the total number of bundle pixels, then scan through each row and until the cumulative number of pixels reaches 50%. We then take the left and rightmost pixel in that row as the starting and ending column. Calibrations are rejected if there is a different number of rows or columns in either the left or the right image.

7.3.3 Resolution and optical power

As a quality check we image a USAF 1951 target in transmission mode with high sampling and for similarity compared the residual and confocal images. Visually there was some misalignment between fibre cores near one edge, however, this did not exceed a distance of 1 core. This can be quantified using the mean of the absolute differences between the images which was 5.5% with no filtering, after applying a Gaussian filter, which is done as part of the imaging process, this is further improved to 0.74%.

There are some optical power losses in the system. The power delivered to the sample was measured as 1 mW. The losses introduced by the beamsplitter configuration alone were 50%, and the intensity is further halved between the two images. The intensity of the line exhibits Gaussian behavior along its elongated axis, to minimize the effects of this the line size is chosen so it overfills the bundle, which leads to a smoothing of intensity accross the line but some power loss from the laser as well as reflections off the metallic connector. Reflections are minimized using absorptive black tape surrounding the bundle.

7.3.4 Mosaicking

Acquired images can be stitched together into a mosaic. The raw images were downsampled to 250x250 pixels. The shift was registered between a template extracted from the current image and the full previous image using normalized cross-correlation [129]. The current image is then inserted dead-leaf into the mosaic, where the pixel values of the new image completely overwrite any previous pixel values, into the mosaic. The bundle diameter is cropped to the central 90% to avoid edge artifacts degrading images. Mosaics and images were generated from recordings. To eliminate the fibre core pattern, especially noticeable with high sampling (Fig. 7.20,7.22), images which were acquired with high sampling were processed with a median filter using a 3x3 pixel region and selecting the central pixel, images acquired with low sampling had inherent smoothing due to the under sampling of the bundle pixels by the camera and were not processed with a median filter.

7.4 Results and discussion

7.4.1 Axial intensity response

System confocal axial response was characterized using axial response profiles. The profiles were generated using a bare fibre moving at known velocity toward a metal plate fluorescently marked with a yellow highlighter. A representative profile is shown in Fig. 7.17. The 3 dB and 10 dB dropoff points occur at the half-width half-max and half-width 10% max of the LS and ELS curves. Signal

drop can be approximated by the relative intensity in the unnormalized ELS and LS profiles at 0 depth.

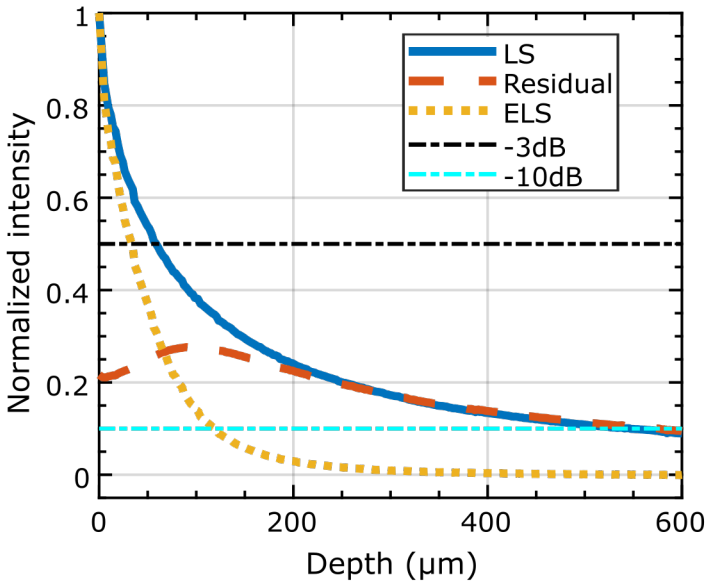


FIGURE 7.17: Representative axial response profile, fibre probe moving towards smooth metal plate. 2.883 μm slit width, 28.883 μm offset. LS and ELS independently normalized, residual shown to the same scale as LS. 3 dB and 10 dB drop off distances for LS and ELS at intersection with respective horizontal lines.

The axial response with a constant slit width is shown in Fig. 7.18(a,b). The LS value was determined by averaging over 7 runs. The 3 and 10 dB dropoff distances were 58 ± 3 and 547 ± 20 μm respectively. The ELS plots show a smaller offset leads to greater improvement in optical sectioning strength for both the 3 dB and 10 dB dropoff positions. As the offset gets larger the 3 dB dropoff distance for ELS converges with the LS much faster than with the 10 dB dropoff distance. Curves with a constant offset and variable slit width are shown in Fig. 7.18(b,c). The slit width improves the 3 dB drop-off for both LS and ELS images, a more narrow slit width improves 10 dB performance in both LS and ELS, however, this happens at a much greater rate for ELS images.

The pixelation of the fibre bundle imposes an additional constraint on sectioning strength: once the offset or the slit width approaches the inter-core spacing of the bundle there are virtually no gains in optical sectioning strength. Three offsets taken near the inter-core spacing of the bundle (at 2.9, 5.8, and 8.6 μm) had 3dB dropoffs at 13.1, 9.6, and 12.5 μm and 10 dB dropoff points of 59.1, 41.9, and 45.1 μm. This indicates virtually no performance gain once the offset approaches the inter-core spacing. The same is true with slit widths near the resolution of the bundle. Three profiles were generated with widths 1.92, 2.9 and 5.8 μm, and with a constant offset of 28.8 μm, the respective LS values were 69.8, 70.4, 66.5 μm, and ELS values were 123.2, 119.2, 119.3 μm indicating there is no further improvement in LS or ELS once the slit width is the smaller than the resolution of the bundle.

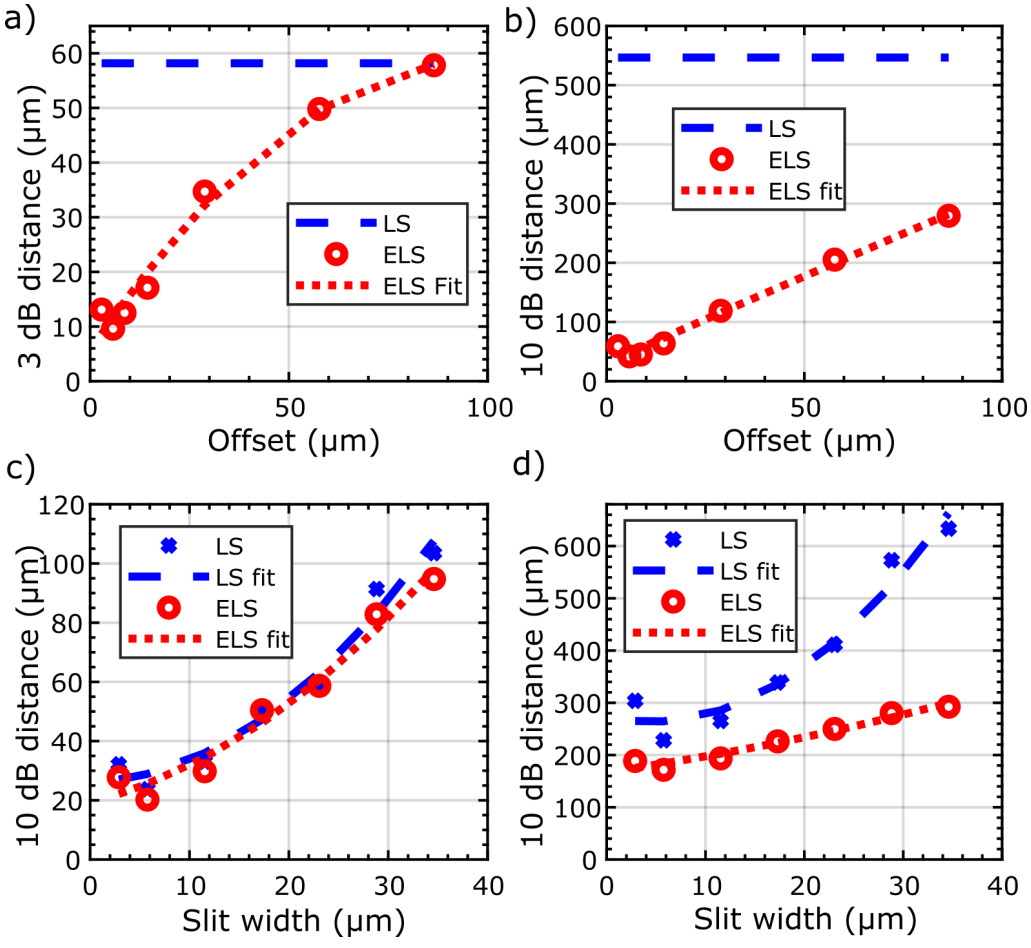


FIGURE 7.18: 3 and 10 dB dropoff distances as a function of offset (a,b) with constant 2.9 μm slit width and as a function of slit width (c,d) with constant 77 μm offset. LS reference dropoff distance in (a) and (b) determined by averaging dropoff distances over 7 runs, mean 3 dB dropoff distance of 58 ± 3 μm and mean 10 dB dropoff distance of 547 ± 18 μm.

7.4.2 Signal drop study

As was shown in the theory (Sect. 7.2.2) with a small offset between illumination and detection slits, in-focus components are subtracted. This leads to a drop in the signal-to-noise ratio. This can come from several sources, including the illumination line extending into the area of the offset slit, cross core coupling of the bundle, and neighboring fluorescence can be excited in the region of the offset slit leading to inaccurate subtraction. Signal drops as a function of slit width and offset are shown in Fig. 7.19. With constant signal width, it is clear the smallest signal drop occurs at the largest offset. At a 77 μm offset, there appears to be virtually no price paid in terms of relative signal drop when using ELS instead of LS. This implies it is possible to obtain the benefits of far-from-focus light rejection with very little penalty paid in signal drop. Although, as will be discussed in Sect. 7.5, a large offset leads to edge artifacts. This signal drop analysis only characterizes the relative signal drop of ELS to LS.

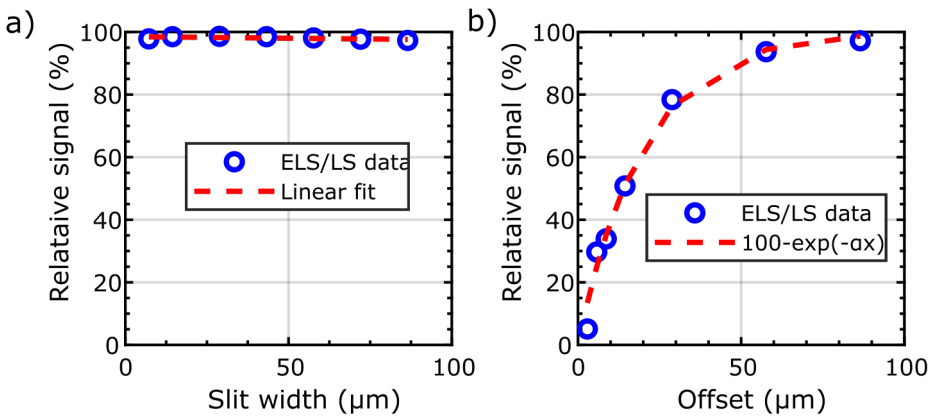


FIGURE 7.19: Percent signal maintained when varying slit width and offset. Relative ELS to LS signal as a function of slit width (a) with constant 77 μm offset, and offset (b) with a constant 2.9 μm slit width, fit determined through non-linear least squares curve fit $\alpha = 0.05$. Signal reported as relative intensity of ELS to LS when bundle is in contact with mirror.

7.4.3 ELSA, ELS+, LS image comparison

This approach offers benefit to a low frame rate system as well as a high speed one. To demonstrate this images of fluorescently stained lens paper over a fluorescent background are acquired at 10 frames/s with a moving probe. Intensity line plots show the improvement in the ELS+ approach. In Fig. 7.20(a), the LS image, the background is caused by far-from-focus light. In Fig. 7.20(b), the ELSA image, the background is caused by a residual image that has shifted and Fig. 7.20(c), the ELS+ image, shows the removal of the out-of-focus background without the interframe artifacts. To normalise the processed images were multiplied iteratively by a scaling factor until 0.01% saturation was reached.

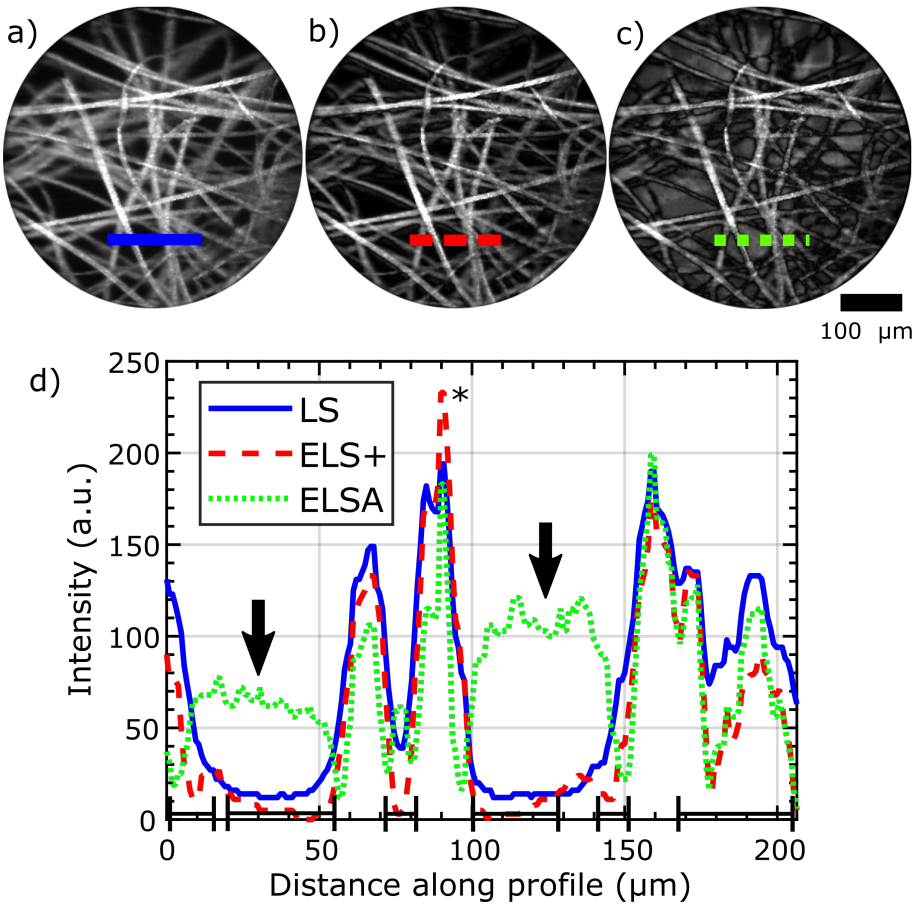


FIGURE 7.20: Lens paper stained with highlighter overlaid on a fluorescent background. (a) LS, (b) ELS+, (c) ELSA, and (d) plot profiles. Images acquired with moving the probe in handheld operation at 10 frames/s. In (a) lower contrast is caused by light returning far-from-focus, while in (c) it results from motion artifacts. Brackets indicate improved sectioning with ELS+, arrows indicate regions where motion artifacts appear in ELSA images. LS and ELS+ were independently normalized to 0.01% saturation which led to the intensity of ELS+ being greater than LS in the marked region (*). ELSA normalized by the same factor as ELS+.

7.4.4 240 frames per second mosaicking

To operate the system at 240 frames/s, the cameras field of view is reduced from 1280 x 960 pixels to 640x480. This is used to take advantage of a higher frame rate. The size of the bundle image needs to be reduced which can be done either by cropping the bundle using a circular aperture in a conjugate plane or by

projecting a less magnified bundle image onto the camera at expense of resolution. The latter was performed. Since only a small change in bundle image size was needed, our approach was to modify the distance between the fibre and objective, as well as tube lens and camera, then finally to adjust the focus of the line by slightly adjusting the cylindrical and collimating lens (L2, L3) positions. An intensity profile taken across a beam indicated uniform symmetric intensity throughout the entire travel of the bundle, while a plot of the trajectory of the beam indicated highly linear motion. The reduced lateral resolution, measured with a USAF target distally illuminated in transmission mode, was 8.8 μm. The 240 frames/s mosaics are shown in Fig. 7.21. Fig. 7.21(a,b) were reconstructed from recording to allow for easy comparison, while Fig. 7.21(c) was acquired with online mosaicking.

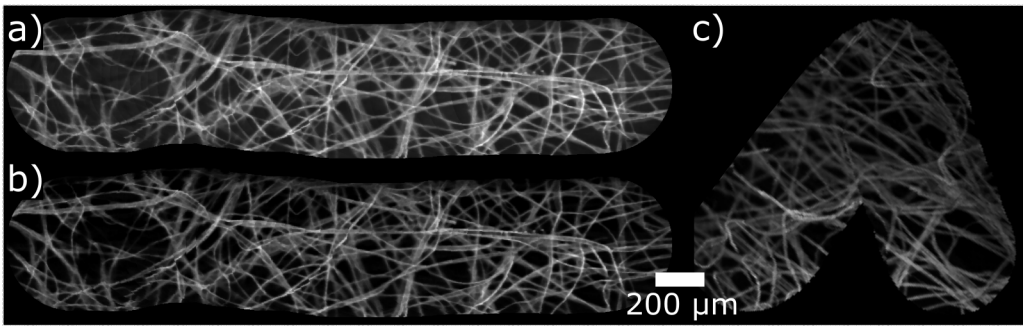


FIGURE 7.21: Lens paper stained with yellow highlighter overlaid on a fluorescent background. Acquired at 240 frames/s. (a) LS, (b) ELS+ reconstructed from recording, (c) ELS+ real-time mosaicing. 8.8 μm resolution.

7.4.5 System validation with porcine oesophagus

We validate this approach for tissue imaging using the *ex vivo* porcine esophagus. The interframe motion artifact reduction can be seen in the single images of Fig. 7.20(b, c) and Fig. 7.22(b, c), where, in both ELSA and ELS+ images the probe had shifted between acquiring the LS image and residual image. In Fig 7.22c, an

ELSA image, constructed from a LS image and residual image, where the bundle had shifted during acquisition is incorrectly subtracted and appears as a cavity. The improvement of ELS over LS can be seen in the mosaics of Fig. 7.20 and Fig. 7.22. While ELS+ approach is robust to interframe artifacts, each line reads out sequentially, and when the probe shifts between lines, intraframe artifacts that occur during the capture of a single frame can still distort reconstructions. These can be further reduced using higher frame rates, which allows for a faster line readout rate.

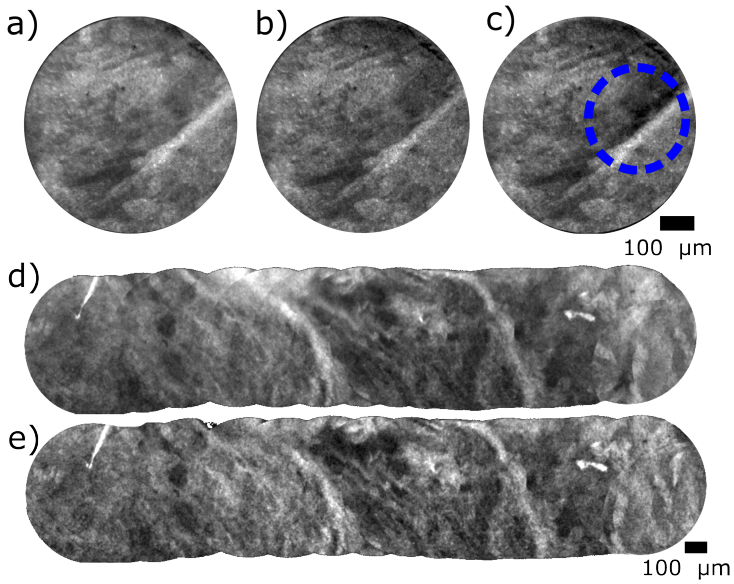


FIGURE 7.22: Porcine oesophagus stained with acriflavine hydrochloride. (a) LS, (b) ELS+, (c) ELSA. Circled region indicates intensity which has been incorrectly subtracted due to a shift between acquisitions. LS and ELS+ independently normalised to 0.01% saturation, ELSA normalised to same factor as ELS+. Mosaics (d) LS, (e) ELS+. Scale bars 100 µm. Acquired in high sampling mode. Acquired at 60 Hz, reconstructed from recording.

7.5 Limitations

Edge artifiacts occur for multiple reasons, some common to both LS and ELS and some specific to ELS. Those that result from using the same fibre for illumination

and detection. Since there is no illumination outside the bundle, out-of-focus light primarily returns to central cores. This means there is inherently better sectioning at the edges. A system with better sectioning performance inherently reduces the effect of this artefact this since more out-of-focus light is rejected.

An offset larger than a few slit widths, such as 76 μm, delivers good performance at the 10 dB dropoff points, however, this comes at the expense of of introducing further artifacts, unique to the ELS approach. These appear in the images acquired with the offset slit and are clipping artifacts and those that result from unequal collection and illumination areas (Fig. 7.23). The clipping artifacts result from the fact that when traversing the bundle face the detection line leads the illumination line, which is clipped by the top of the bundle in the residual image. Once the illumination line becomes incident on the bundle it introduces an additional problem since the relative areas of the illumination line to the detection slit are not uniform during scanning. When the offset images are subtracted from LS images, there is no improved sectioning in clipped areas of the bundle.

The simplest way to reduce these artefacts is to minimize the offset. This results in better sectioning, but also a lower signal to noise ratio. Fig. 7.23 shows this very clearly. When an offset of 11.53 μm is used there are virtually no artifacts from clipping or unequal illumination areas. A flat field correction can also be used, but this does not correct clipping artifacts. While it may be tempting to select a large offset to attenuate only the far-from-focus light and keep signal high there is a high price to be paid in terms of clipping artifacts.

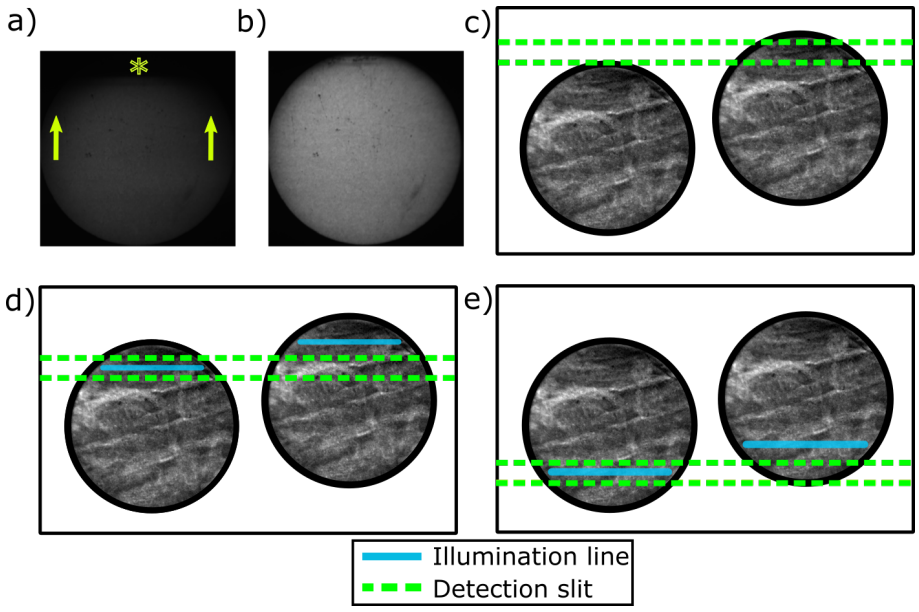


FIGURE 7.23: Average of ~300 residual images acquired using a constant 8.7 μm slit width with (a) 74 μm offset and (b) 12 μm offset. Images were taken with a moving probe over paper stained with a yellow highlighter. Artifacts identified in (a) are those caused by clipping (⋆) and differences in collection and illumination areas of the bundle (arrows). Experimental causes of artifacts are: (c) illumination line clipping, (d) greater detection area in residual images, (e) smaller detection area in residual images.

7.6 Conclusion

The newly reported approach of simultaneous acquisition enhanced LS restores the frame rate of a LS only system without inter-frame motion artifacts. We have validated the system with porcine esophagus and shown a reduction in features caused by incorrectly subtracted light due to motion artefacts in non-simultaneous ELS approaches. The system reported here, with a frame rate of 240 Hz, is faster than typical research systems and an order of magnitude faster than currently available commercial systems.

Chapter 8

Conclusion and future work

Fibre bundle endomicroscopy has emerged to enable real time imaging of cellular level features *in-vivo*. Its usefulness as an imaging modality lies in its ability to give clinicians cellular level information with good feature discrimination. It is also useful to generate mosaics to create a larger field-of-view without a larger and more invasive probe.

The SIM section presented a clinically viable DMD based endomicroscope for improved contrast. Widefield images acquired with systems like the HRME, when imaging thick tissue are degraded by out-of-focus light. A parameter set and images clearly showing improved sectioning were demonstrated. The on-the-fly tunability of the modulation frequency of a DMD endomicroscope means that sectioning strength and signal can be quickly balanced by clinicians. Tissue imaging was demonstrated using stained bovine stomach. The frame rate was demonstrated at a processed 40 frames/s, around 3-4 times faster than current commercial confocal laser endomicroscopy systems.

The on-the-fly modulation tunability of the DMD based system ought to enable the depth sectioning in real-time. The lateral resolution does vary with depth and depth slices won't have identical lateral resoltuions. One could envisage an

approach which could expand the depth-of-field such as through numerical aperture confinement [89] to reduce the effect of this.

The Tetris and Four Quad approaches delivered equivalent contrast transfer to SIM when using spatial frequencies that were 3x higher. This led to similar sectioning and reduced artefacts with with Tetris this, however, improvement came at the expense of 40 patterns vs the 3 for SIM, making it impractical. Four Quad demonstrated improved sectioning over SIM but with significantly reduced artefacts caused by the halved-core problem. The resolution to this for Tetris was pattern shifting. Future work to exploit the improved contrast transfer involves developing a Four Quad detection scheme to overcome the halved core problem using a reasonable amount of patterns.

SIM allows for low cost optical sectioning, at a relatively high frame rate, motion between frames has led to artefacts, making the technique difficult to use *in-vivo* and preventing the use of mosaicking to synthesize a larger effective field of view. The automatic motion compensation technique, by reorienting line patterns to the direction of motion will allow a DMD based device to be used in a clinical setting. Advancing the usefulness of low-cost devices with optical sectioning and the increased field-of-view enabled by mosaicking. Future work may involve combining this with depth sectioning for multi-depth mosaics.

The final development presented was an approach to improve the motion artefacts introduced in a subtraction-based enhanced line scanning (ELS) endomicroscope. Previous subtraction based enhanced line scan systems relied on either two frames, or used two cameras, the approach presented here allowed for real time subtraction in a single frame, eliminating the artefacts that occur between

frames. The novel approach made way for real-time mosaicing at 240 frames/s.

Bibliography

[1] B. D. Grant, J. H. Fregnani, J. C. Possati Resende, C. Scapulatempo-Neto, G. M. Matsushita, E. C. Mauad, T. Quang, M. H. Stoler, P. E. Castle, K. M. Schmeler, and R. R. Richards-Kortum, "High-resolution microendoscopy: A point-of-care diagnostic for cervical dysplasia in low-resource settings", *Eur J Cancer Prev*, vol. 26, no. 1, pp. 63–70, 2017, ISSN: 1473-5709 (Electronic) 0959-8278 (Linking). DOI: 10.1097/CEJ.0000000000000219. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/26637074.

[2] J. M. Jabbour, M. A. Saldua, J. N. Bixler, and K. C. Maitland, "Confocal endomicroscopy: Instrumentation and medical applications", *Ann Biomed Eng*, vol. 40, no. 2, pp. 378–97, 2012, ISSN: 1573-9686 (Electronic) 0090-6964 (Linking). DOI: 10.1007/s10439-011-0426-y. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/21994069.

[3] K. B. Sung, R. Richards-Kortum, M. Follen, A. Malpica, C. Liang, and M. Descour, "Fiber optic confocal reflectance microscopy: A new real-time technique to view nuclear morphology in cervical squamous epithelium in vivo", *Opt Express*, vol. 11, no. 24, pp. 3171–81, 2003, ISSN: 1094-4087 (Electronic) 1094-4087 (Linking). DOI: 10.1364/oe.11.003171. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/19471442.

[4] J. Tan, M. Quinn, J. Pyman, P. Delaney, and W. McLaren, "Detection of cervical intraepithelial neoplasia in vivo using confocal endomicroscopy", *BJOG: An International Journal of Obstetrics and Gynaecology*, vol. 116, no. 12, pp. 1663–1670, 2009, ISSN: 1470-0328. DOI: 10.1111/j.1471-0528.2009.02261.x. [Online]. Available: https://dx.doi.org/10.1111/j.1471-0528.2009.02261.x.

[5] T. Wilson and C. Sheppard, "Theory and practice of scanning optical microscopy", *London Academic Press*, vol. 1, 1984.

[6] F. Zernike, "How i discovered phase contrast", *Science*, vol. 121, no. 3141, pp. 345–349, 1955. DOI: 10.1126/science.121.3141.345. [Online]. Available: https://science.sciencemag.org/content/sci/121/3141/345.full.pdf.

[7] D. Axelrod, "Zero-cost modification of bright field microscopes for imaging phase gradient on cells: Schlieren optics", *Cell Biophys*, vol. 3, no. 2, pp. 167–73, 1981, ISSN: 0163-4992 (Print) 0163-4992 (Linking). DOI: 10.1007/BF02788132. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/6168379.

[8] J. Li, T. G. Bifano, and J. Mertz, "Widefield fluorescence microscopy with sensor-based conjugate adaptive optics using oblique back illumination", *J Biomed Opt*, vol. 21, no. 12, p. 121 504, 2016, ISSN: 1560-2281 (Electronic) 1083-3668 (Linking). DOI: 10.1117/1.JBO.21.12.121504. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/27653793.

[9] C. Longo, M. Ragazzi, M. Rajadhyaksha, K. Nehal, A. Bennassar, G. Pellacani, and J. Malvehy Guilera, "In vivo and ex vivo confocal microscopy for dermatologic and mohs surgeons", *Dermatol Clin*, vol. 34, no. 4, pp. 497–504, 2016, ISSN: 1558-0520 (Electronic) 0733-8635 (Linking). DOI: 10.1016/j.det.2016.05.012. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/27692455.

[10] Albani, "Structure and dynamics of macromolecules: Absorption and fluorescence studies", 2004.

[11] P. K. Miller, R. K. Roenigk, D. G. Brodland, and H. W. Randle, "Cutaneous micrographic surgery: Mohs procedure", *Mayo Clinic Proceedings*, vol. 67, no. 10, pp. 971–980, 1992, ISSN: 00256196. DOI: 10.1016/s0025-6196(12)60929-x.

[12] C. Q. Li, T. Yu, X. L. Zuo, X. J. Xie, W. B. Li, C. L. Chu, F. Zuo, and Y. Q. Li, "Effects on confocal laser endomicroscopy image quality by different acriflavine concentrations", *J Interv Gastroenterol*, vol. 1, no. 2, pp. 59–63, 2011, ISSN: 2154-1280 (Print) 2154-1280 (Linking). DOI: 10.4161/jig.1.2.16828. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/21776427.

[13] M. Minsky, "Confocal patent", *US Patent Office*, 1961.

[14] T. Vercauteren, A. Meining, F. Lacombe, and A. Perchant, *Real time autonomous video image registration for endomicroscopy: fighting the compromises*, ser. SPIE BiOS. SPIE, 2008, vol. 6861. [Online]. Available: https://doi.org/10.1117/12.763089.

[15] T. Vercauteren, A. Perchant, G. Malandain, X. Pennec, and N. Ayache, "Robust mosaicing with correction of motion distortions and tissue deformations for in vivo fibered microscopy", *Med Image Anal*, vol. 10, no. 5,

pp. 673–92, 2006, ISSN: 1361-8415 (Print) 1361-8415 (Linking). DOI: 10.1016/j.media.2006.06.006. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/16887375.

[16] T. Vercauteren, A. Perchant, X. Pennec, and N. Ayache, "Mosaicing of confocal microscopic in vivo soft tissue video sequences", *Med Image Comput Comput Assist Interv*, vol. 8, no. Pt 1, pp. 753–60, 2005. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/16685914.

[17] Conference Paper, 2019. DOI: 10.1117/12.2509590.

[18] M. Goetz, T. Toermer, M. Vieth, K. Dunbar, A. Hoffman, P. R. Galle, M. F. Neurath, P. Delaney, and R. Kiesslich, "Simultaneous confocal laser endomicroscopy and chromoendoscopy with topical cresyl violet", *Gastrointestinal endoscopy*, vol. 70, no. 5, pp. 959–968, 2009, ISSN: 0016-5107.

[19] P. S. Thong, M. C. Olivo, R. Bhuvaneswari, S. S. Tandjung, M. M. Movania, W.-M. Chiew, H.-S. Seah, F. Lin, K. Qian, and K.-C. Soo, "Toward real-time virtual biopsy of oral lesions using confocal laser endomicroscopy interfaced with embedded computing", *Journal of biomedical optics*, vol. 17, no. 5, p. 056 009, 2012, ISSN: 1083-3668.

[20] M. Degueldre, J. Vandromme, A. de Wind, and F. Feoli, "Real-time in-vivo microscopic imaging of the cervix using confocal laser endomicroscopy: Preliminary observations and feasibility study", *European Journal of Cancer Prevention*, vol. 25, no. 4, pp. 335–343, 2016, ISSN: 0959-8278.

[21] C. Trovato, A. Sonzogni, D. Ravizza, G. Fiori, D. Tamayo, G. De Roberto, A. de Leone, S. De Lisi, and C. Crosta, "Confocal laser endomicroscopy for in vivo diagnosis of barrett's oesophagus and associated neoplasia: A pilot study conducted in a single italian centre", *Digestive and Liver Disease*, vol. 45, no. 5, pp. 396–402, 2013, ISSN: 1590-8658.

[22] A. M. Buchner, M. W. Shahid, M. G. Heckman, M. Krishna, M. Ghabril, M. Hasan, J. E. Crook, V. Gomez, M. Raimondo, and T. Woodward, "Comparison of probe-based confocal laser endomicroscopy with virtual chromoendoscopy for classification of colon polyps", *Gastroenterology*, vol. 138, no. 3, pp. 834–842, 2010, ISSN: 0016-5085.

[23] L. Thiberville, M. Salaün, S. Lachkar, S. Dominique, S. Moreno-Swirc, C. Vever-Bizet, and G. Bourg-Heckly, "Confocal fluorescence endomicroscopy of the human airways", *Proceedings of the American Thoracic Society*, vol. 6, no. 5, pp. 444–449, 2009, ISSN: 1546-3222.

[24] N. L. Martirosyan, J. Georges, J. M. Eschbacher, D. D. Cavalcanti, A. M. Elhadi, M. G. Abdelwahab, A. C. Scheck, P. Nakaji, R. F. Spetzler, and M. C. Preul, "Potential application of a handheld confocal endomicroscope imaging system using a variety of fluorophores in experimental gliomas and normal brain", *Neurosurgical focus*, vol. 36, no. 2, E16, 2014, ISSN: 1092-0684.

[25] D. Panarello, E. Compérat, O. Seyde, A. Colau, C. Terrone, and B. Guillonneau, "Atlas of ex vivo prostate tissue and cancer images using confocal laser endomicroscopy: A project for intraoperative positive surgical margin detection during radical prostatectomy", *European urology focus*, vol. 6, no. 5, pp. 941–958, 2020, ISSN: 2405-4569.

[26] V. Becker, M. B. Wallace, P. Fockens, S. von Delius, T. A. Woodward, M. Raimondo, R. P. Voermans, and A. Meining, "Needle-based confocal endomicroscopy for in vivo histology of intra-abdominal organs: First results in a porcine model (with)", *Gastrointestinal endoscopy*, vol. 71, no. 7, pp. 1260–1266, 2010, ISSN: 0016-5107.

[27] S. Zuo, M. Hughes, and G.-Z. Yang, "Novel balloon surface scanning device for intraoperative breast endomicroscopy", *Annals of biomedical engineering*, vol. 44, no. 7, pp. 2313–2326, 2016, ISSN: 0090-6964.

[28] A. L. Polglase, W. J. McLaren, S. A. Skinner, R. Kiesslich, M. F. Neurath, and P. M. Delaney, "A fluorescence confocal endomicroscope for in vivo microscopy of the upper- and the lower-gi tract", *Gastrointest Endosc*, vol. 62, no. 5, pp. 686–95, 2005, ISSN: 0016-5107 (Print) 0016-5107 (Linking). DOI: 10.1016/j.gie.2005.05.021. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/16246680.

[29] I. Odagi, T. Kato, H. Imazu, M. Kaise, S. Omar, and H. Tajiri, "Examination of normal intestine using confocal endomicroscopy", *Journal of gastroenterology and hepatology*, vol. 22, no. 5, pp. 658–662, 2007, ISSN: 0815-9319.

[30] T. P. Chang, D. R. Leff, S. Shousha, D. J. Hadjiminas, R. Ramakrishnan, M. R. Hughes, G. Z. Yang, and A. Darzi, "Imaging breast cancer morphology using probe-based confocal laser endomicroscopy: Towards a real-time intraoperative imaging tool for cavity scanning", *Breast Cancer Res Treat*, vol. 153, no. 2, pp. 299–310, 2015, ISSN: 1573-7217 (Electronic) 0167-6806 (Linking). DOI: 10.1007/s10549-015-3543-8. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/26283299.

[31] G. D. De Palma, D. Esposito, G. Luglio, G. Limite, A. Accurso, V. Sollazzo, F. Maione, G. Cassese, S. Siciliano, N. Gennarelli, G. Ilardi, M. Paternoster, M. C. Giglio, and P. Forestieri, "Confocal laser endomicroscopy in breast surgery: A pilot study", *BMC Cancer*, vol. 15, no. 1, 2015, ISSN: 1471-2407. DOI: 10.1186/s12885-015-1245-6. [Online]. Available: https://dx.doi.org/10.1186/s12885-015-1245-6.

[32] J. T. Liu, D. Meza, and N. Sanai, "Trends in fluorescence image-guided surgery for gliomas", *Neurosurgery*, vol. 75, no. 1, pp. 61–71, 2014, ISSN: 1524-4040 (Electronic) 0148-396X (Linking). DOI: 10.1227/NEU.0000000000000034 [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/24618801.

[33] E. Belykh, E. J. Miller, A. Carotenuto, A. A. Patel, C. Cavallo, N. L. Martirosyan, D. R. Healey, V. A. Byvaltsev, A. C. Scheck, M. T. Lawton, J. M. Eschbacher, P. Nakaji, and M. C. Preul, "Progress in confocal laser endomicroscopy for neurosurgery and technical nuances for brain tumor imaging with fluorescein", *Front Oncol*, vol. 9, p. 554, 2019, ISSN: 2234-943X (Print) 2234-943X (Linking). DOI: 10.3389/fonc.2019.00554. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/31334106.

[34] I. Wolf, J. Smolle, H. Soyer, and H. Kerl, "Sensitivity in the clinical diagnosis of malignant melanoma", *Melanoma Research*, vol. 8, pp. 425–429, 1998.

[35] A. Gerger, S. Koller, T. Kern, C. Massone, K. Steiger, E. Richtig, H. Kerl, and J. Smolle, "Diagnostic applicability of in vivo confocal laser scanning microscopy in melanocytic skin tumors", *J Invest Dermatol*, vol. 124, no. 3, pp. 493–8, 2005, ISSN: 0022-202X (Print) 0022-202X (Linking). DOI: 10.1111/j.0022-202X.2004.23569.x. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/15737188.

[36] G. Ball, J. Demmerle, R. Kaufmann, I. Davis, I. M. Dobbie, and L. Schermelleh, "Simcheck: A toolbox for successful super-resolution structured illumination microscopy", *Sci Rep*, vol. 5, p. 15915, 2015, ISSN: 2045-2322 (Electronic) 2045-2322 (Linking). DOI: 10.1038/srep15915. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/26525406.

[37] D. Shin, M. A. Protano, A. D. Polydorides, S. M. Dawsey, M. C. Pierce, M. K. Kim, R. A. Schwarz, T. Quang, N. Parikh, M. S. Bhutani, F. Zhang, G. Wang, L. Xue, X. Wang, H. Xu, S. Anandasabapathy, and R. R. Richards-Kortum, "Quantitative analysis of high-resolution microendoscopic images for diagnosis of esophageal squamous cell carcinoma", *Clin Gastroenterol Hepatol*, vol. 13, no. 2, 272–279 e2, 2015, ISSN: 1542-7714 (Electronic)

1542-3565 (Linking). DOI: 10.1016/j.cgh.2014.07.030. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/25066838.

[38] M. C. Pierce, Y. Guan, M. K. Quinn, X. Zhang, W. H. Zhang, Y. L. Qiao, P. Castle, and R. Richards-Kortum, "A pilot study of low-cost, high-resolution microendoscopy as a tool for identifying women with cervical precancer", *Cancer Prev Res (Phila)*, vol. 5, no. 11, pp. 1273–9, 2012, ISSN: 1940-6215 (Electronic) 1940-6215 (Linking). DOI: 10.1158/1940-6207.CAPR-12-0221. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/22926339.

[39] V. Kayamba, A. Shibemba, K. Zyambo, D. C. Heimburger, D. R. Morgan, and P. Kelly, "High prevalence of gastric intestinal metaplasia detected by confocal laser endomicroscopy in zambian adults", *PloS one*, vol. 12, no. 9, e0184272, 2017, ISSN: 1932-6203.

[40] E. C. Yang, I. S. Vohra, H. Badaoui, R. A. Schwarz, K. D. Cherry, J. Jacob, J. Rodriguez, M. D. Williams, N. Vigneswaran, A. M. Gillenwater, and R. R. Richards-Kortum, "Prospective evaluation of oral premalignant lesions using a multimodal imaging system: A pilot study", *Head Neck*, vol. 42, no. 2, pp. 171–179, 2020, ISSN: 1097-0347 (Electronic) 1043-3074 (Linking). DOI: 10.1002/hed.25978. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/31621979.

[41] D. B. Ellebrecht, S. Latus, A. Schlaefer, T. Keck, and N. Gessert, "Towards an optical biopsy during visceral surgical interventions", *Visc Med*, vol. 36, no. 2, pp. 70–79, 2020, ISSN: 2297-4725 (Print) 2297-4725 (Linking). DOI: 10.1159/000505938. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/32355663.

[42] P. A. Keahey, T. S. Tkaczyk, K. M. Schmeler, and R. R. Richards-Kortum, "Optimizing modulation frequency for structured illumination in a fiber-optic microendoscope to image nuclear morphometry in columnar epithelium", *Biomed Opt Express*, vol. 6, no. 3, pp. 870–80, 2015, ISSN: 2156-7085 (Print) 2156-7085 (Linking). DOI: 10.1364/BOE.6.000870. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/25798311.

[43] N. Bozinovic, C. Ventalon, T. Ford, and J. Mertz, "Fluorescence endomicroscopy with structured illumination", *Opt Express*, vol. 16, no. 11, pp. 8016–25, 2008, ISSN: 1094-4087 (Electronic) 1094-4087 (Linking). DOI: 10.1364/oe.16.008016. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/18545511.

[44] C. Ventalon, J. Mertz, and V. Emilani, "Depth encoding with lensless structured illumination fluorescence micro-endoscopy", 2015.

[45] M. J. Marques, M. R. Hughes, K. Vyas, A. Thrapp, H. Zhang, A. Bradu, G. Gelikonov, P. Giataganas, C. J. Payne, G. Z. Yang, and A. Podoleanu, "En-face optical coherence tomography/fluorescence endomicroscopy for minimally invasive imaging using a robotic scanner", *J Biomed Opt*, vol. 24, no. 6, pp. 1–15, 2019, ISSN: 1560-2281 (Electronic) 1083-3668 (Linking). DOI: 10.1117/1.JBO.24.6.066006. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/31222989.

[46] A. Thrapp and M. Hughes, "Automatic motion compensation for structured illumination endomicroscopy using a flexible fiber bundle", *J Biomed Opt*, vol. 25, no. 2, pp. 1–13, 2020, ISSN: 1560-2281 (Electronic) 1083-3668 (Linking). DOI: 10.1117/1.JBO.25.2.026501. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/32100492.

[47] M. Pierce, D. Yu, and R. Richards-Kortum, "High-resolution fiber-optic microendoscopy for in situ cellular imaging", *J Vis Exp*, no. 47, 2011, ISSN: 1940-087X (Electronic) 1940-087X (Linking). DOI: 10.3791/2306. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/21248707.

[48] T. J. Muldoon, M. C. Pierce, D. L. Nida, M. D. Williams, A. Gillenwater, and R. Richards-Kortum, "Subcellular-resolution molecular imaging within living tissue by fiber microendoscopy", *Opt Express*, vol. 15, no. 25, pp. 16 413–23, 2007, ISSN: 1094-4087 (Electronic) 1094-4087 (Linking). DOI: 10.1364/oe.15.016413. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/19550931.

[49] X. Chen, K. L. Reichenbach, and C. Xu, "Experimental and theoretical analysis of core-to-core coupling on fiber bundle imaging", *Opt Express*, vol. 16, no. 26, pp. 21 598–607, 2008, ISSN: 1094-4087 (Electronic) 1094-4087 (Linking). DOI: 10.1364/oe.16.021598. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/19104591.

[50] H. E. Parker, A. Perperidis, A. Karam-Eldaly, Y. Altmann, K. Dhaliwal, R. R. Thomson, S. McLaughlin, and M. G. Tanner, "Characterising cross-coupling in coherent fibre bundles", in *European Conference on Biomedical Optics*, Optical Society of America, 11076_50.

[51] A. W. Snyder and J. D. Love, "Optical waveguide theory", *J. Opt. Soc. Amer. A*, vol. 3, no. 3, pp. 378–379, 1986.

[52] Generic, 2019. [Online]. Available: https://nanohub.org/resources/29411.

[53] A. Perperidis, H. E. Parker, A. Karam-Eldaly, Y. Altmann, K. Dhaliwal, R. R. Thomson, M. G. Tanner, and S. McLaughlin, "Characterization and modelling of inter-core coupling in coherent fiber bundles", *Optics Express*, vol. 25, no. 10, pp. 11 932–11 953, 2017, ISSN: 1094-4087.

[54] K. L. Reichenbach and C. Xu, "Numerical analysis of light propagation in image fibers or coherent fiber bundles", *Optics express*, vol. 15, no. 5, pp. 2151–2165, 2007, ISSN: 1094-4087.

[55] G. S. D. Gordon, J. Joseph, T. Sawyer, A. J. Macfaden, C. Williams, T. D. Wilkinson, and S. E. Bohndiek, "Full-field quantitative phase and polarisation-resolved imaging through an optical fibre bundle", *Opt Express*, vol. 27, no. 17, pp. 23 929–23 947, 2019, ISSN: 1094-4087 (Electronic) 1094-4087 (Linking). DOI: 10.1364/OE.27.023929. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/31510290.

[56] R. B. Dyott, J. Cozens, and D. Morris, "Preservation of polarisation in optical-fibre waveguides with elliptical cores", *Electronics Letters*, vol. 15, no. 13, pp. 380–382, 1979, ISSN: 1350-911X.

[57] A. K. Eldaly, Y. Altmann, A. Perperidis, N. Krstajić, T. R. Choudhary, K. Dhaliwal, and S. McLaughlin, "Deconvolution and restoration of optical endomicroscopy images", *IEEE Transactions on Computational Imaging*, vol. 4, no. 2, pp. 194–205, 2018, ISSN: 2333-9403.

[58] M. Hughes, P. Giataganas, and G.-Z. Yang, "Color reflectance fiber bundle endomicroscopy without back-reflections", *Journal of Biomedical Optics*, vol. 19, no. 3, p. 030 501, 2014. [Online]. Available: https://doi.org/10.1117/1.JBO.19.3.030501.

[59] R. Juškattis, T. Wilson, and T. Watson, "Real-time white light reflection confocal microscopy using a fibre-optic bundle", *Scanning*, vol. 19, no. 1, pp. 15–19, 1997, ISSN: 0161-0457.

[60] M. Hughes, T. P. Chang, and G. Z. Yang, "Fiber bundle endocytoscopy", *Biomed Opt Express*, vol. 4, no. 12, pp. 2781–94, 2013, ISSN: 2156-7085 (Print) 2156-7085 (Linking). DOI: 10.1364/BOE.4.002781. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/24409380.

[61] J. C. Mertz, T. N. Ford, and K. K. Chu, "Phase contrast microscopy with oblique back-illumination", *US Patent Office*, 2015.

[62] J. Mertz, A. Gasecka, A. Daradich, I. Davison, and D. Coté, "Phase-gradient contrast in thick tissue with a scanning microscope", *Biomedical Optics Express*, vol. 5, no. 2, p. 407, 2014, ISSN: 2156-7085. DOI: 10.1364/boe.5.000407. [Online]. Available: https://dx.doi.org/10.1364/boe.5.000407%20https://kp-pdf.s3.amazonaws.com/d90dc1e8-09eb-43ab-bc94-a9abe870677c.pdf?X-Amz-Algorithm=AWS4-HMAC-SHA256&X-Amz-Credential=AKIAUROH2NUQSIQZIEG4%2F20200605%2Fus-east-1%2Fs3%2Faws4_request&X-Amz-Date=20200605T153306Z&X-Amz-Expires=600&X-Amz-SignedHeaders=host&X-Amz-Signature=a5a3bdcd527cf4f99b1a34c6a299d00ebb68440af

[63] M. Kyrish, J. Dobbs, S. Jain, X. Wang, D. Yu, R. Richards-Kortum, and T. S. Tkaczyk, "Needle-based fluorescence endomicroscopy via structured illumination with a plastic, achromatic objective", *J Biomed Opt*, vol. 18, no. 9, p. 096 003, 2013, ISSN: 1560-2281 (Electronic) 1083-3668 (Linking). DOI: 10.1117/1.JBO.18.9.096003. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/24002190.

[64] "Cellvizio product brochure", 2020. [Online]. Available: https://www.maunakeatech.com/en/cellvizio.

[65] D. Jacobs, "Class notes on correlation and convolution", 2005.

[66] J. Jonkman and C. M. Brown, "Any way you slice it-a comparison of confocal microscopy techniques", *J Biomol Tech*, vol. 26, no. 2, pp. 54–65, 2015, ISSN: 1943-4731 (Electronic) 1524-0215 (Linking). DOI: 10.7171/jbt.15-2602-003. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/25802490.

[67] T. Wilson and A. R. Carlini, "Size of the detector in confocal imaging systems", *Optics Letters*, vol. 12, no. 4, pp. 227–229, 1987. DOI: 10.1364/OL.12.000227. [Online]. Available: http://ol.osa.org/abstract.cfm?URI=ol-12-4-227.

[68] C. J. R. Sheppard, M. Gu, and M. Roy, "Signal-to-noise ratio in confocal microscope systems", *Journal of Microscopy*, vol. 168, no. 3, pp. 209–218, 1992, ISSN: 0022-2720. DOI: 10.1111/j.1365-2818.1992.tb03264.x. [Online]. Available: https://doi.org/10.1111/j.1365-2818.1992.tb03264.x.

[69] D. Sandison, D. Piston, R. Williams, and W. Webb, "Quantitative comparison of background rejection, signal-to-noise ratio, and resolution in confocal and full-field laser scanning microscopes", *Applied optics*, vol. 34, no. 19, pp. 3576–3588, 1995, ISSN: 2155-3165.

[70] M. Dabbs T Fau - Glass and M. Glass, "Fiber-optic confocal microscope: Focon", no. 1559-128X (Print), 1992.

[71] M. R. Harris, "Scanning confocal microscope including a single fibre for transmitting light to and receiving light from an object", 1992.

[72] L. Giniunas, R. Juskaitis, and S. V. Shatalin, "Scanning fibre-optic microscope", *Electronics Letters*, vol. 27, no. 9, pp. 724–726, 1991, ISSN: 0013-5194. DOI: 10.1049/el:19910450.

[73] P. M. Delaney, M. R. Harris, and R. G. King, "Fiber-optic laser scanning confocal microscope suitable for fluorescence imaging", *Applied Optics*, vol. 33, no. 4, p. 573, 1994, ISSN: 0003-6935. DOI: 10.1364/ao.33.000573. [Online]. Available: https://dx.doi.org/10.1364/AO.33.000573%20https://kp-pdf.s3.amazonaws.com/e1c99252-b43d-4ec1-9365-21fdf30e89a6.pdf?X-Amz-Algorithm=AWS4-HMAC-SHA256&X-Amz-Credential=AKIAUROH2NUQSIQZIEG4%2F20200608%2Fus-east-1%2Fs3%2Faws4_request&X-Amz-Date=20200608T131508Z&X-Amz-Expires=600&X-Amz-SignedHeaders=host&X-Amz-Signature=0eaf9dc3a719e8e1af112a9e5a74d3b1e02e34030

[74] P. Demétrio de Souza França, N. Guru, S. Roberts, S. Kossatz, C. Mason, M. Abrahão, R. A. Ghossein, S. G. Patel, and T. Reiner, "Fluorescence-guided resection of tumors in mouse models of oral cancer", *Scientific Reports*, vol. 10, no. 1, p. 11 175, 2020, ISSN: 2045-2322. DOI: 10.1038/s41598-020-67958-8. [Online]. Available: https://doi.org/10.1038/s41598-020-67958-8.

[75] A. F. Gmitro and D. Aziz, "Confocal microscopy through a fiber-optic imaging bundle", *Opt Lett*, vol. 18, no. 8, p. 565, 1993, ISSN: 0146-9592 (Print) 0146-9592 (Linking). DOI: 10.1364/ol.18.000565. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/19802201.

[76] P. M. Lane, A. L. P. Dlugan, R. Richards-Kortum, and C. E. MacAulay, "Fiber-optic confocal microscopy using a spatial light modulator", *Optics Letters*, vol. 25, no. 24, pp. 1780–1782, 2000, ISSN: 0146-9592. DOI: Doi10.1364/Ol.25.001780. [Online]. Available: %3CGo%20to%20ISI%3E://WOS:000165814600012.

[77] Q. Wang, J. Zheng, K. Wang, K. Gui, H. Guo, and S. Zhuang, "Parallel detection experiment of fluorescence confocal microscopy using dmd", *Scanning*, vol. 38, no. 3, pp. 234–239, 2016, ISSN: 0161-0457. DOI: 10.1002/sca.21265. [Online]. Available: https://dx.doi.org/10.1002/sca.21265.

[78] M. Hughes and G. Z. Yang, "High speed, line-scanning, fiber bundle fluorescence confocal endomicroscopy for improved mosaicking", *Biomed Opt Express*, vol. 6, no. 4, pp. 1241–52, 2015, ISSN: 2156-7085 (Print) 2156-7085 (Linking). DOI: 10.1364/BOE.6.001241. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/25909008.

[79] ——, "Line-scanning fiber bundle endomicroscopy with a virtual detector slit", *Biomed Opt Express*, vol. 7, no. 6, pp. 2257–68, 2016, ISSN: 2156-7085 (Print) 2156-7085 (Linking). DOI: 10.1364/BOE.7.002257. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/27375942.

[80] Y. Tang, J. Carns, and R. R. Richards-Kortum, "Line-scanning confocal microendoscope for nuclear morphometry imaging", *J Biomed Opt*, vol. 22, no. 11, pp. 1–6, 2017, ISSN: 1560-2281 (Electronic) 1083-3668 (Linking). DOI: 10.1117/1.JBO.22.11.116005. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/29129041.

[81] Y. Tang, A. Kortum, I. Vohra, M. Othman, S. Dhingra, N. Mansour, J. Carns, S. Anandasabapathy, and R. Richards-Kortum, "Improving nuclear morphometry imaging with real-time and low-cost line-scanning confocal microendoscope", *Opt Lett*, vol. 44, no. 3, pp. 654–657, 2019, ISSN: 1539-4794 (Electronic) 0146-9592 (Linking). DOI: 10.1364/OL.44.000654. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/30702702.

[82] V. Poher, G. T. Kennedy, H. B. Manning, D. M. Owen, H. X. Zhang, E. Gu, M. D. Dawson, P. M. French, and M. A. Neil, "Improved sectioning in a slit scanning confocal microscope", *Opt Lett*, vol. 33, no. 16, pp. 1813–5, 2008, ISSN: 0146-9592 (Print) 0146-9592 (Linking). DOI: 10.1364/ol.33.001813. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/18709096.

[83] Y. Tang, A. Kortum, I. Vohra, J. Carns, S. Anandasabapathy, and R. Richards-Kortum, "Simple differential digital confocal aperture to improve axial response of line-scanning confocal microendoscopes", *Opt Lett*, vol. 44, no. 18, pp. 4519–4522, 2019, ISSN: 1539-4794 (Electronic) 0146-9592 (Linking). DOI: 10.1364/OL.44.004519. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/31517920.

[84] P. Kim, E. Chung, H. Yamashita, K. E. Hung, A. Mizoguchi, R. Kucherlapati, D. Fukumura, R. K. Jain, and S. H. Yun, "In vivo wide-area cellular imaging by side-view endomicroscopy", *Nat Methods*, vol. 7, no. 4, pp. 303–5, 2010, ISSN: 1548-7105 (Electronic) 1548-7091 (Linking). DOI: 10.

1038/nmeth.1440. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/20228814.

[85] X. Duan, H. Li, J. Zhou, Q. Zhou, K. R. Oldham, and T. D. Wang, "Visualizing epithelial expression of egfr in vivo with distal scanning side-viewing confocal endomicroscope", *Sci Rep*, vol. 6, p. 37 315, 2016, ISSN: 2045-2322 (Electronic) 2045-2322 (Linking). DOI: 10.1038/srep37315. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/27874037.

[86] T. N. Ford, D. Lim, and J. Mertz, "Fast optically sectioned fluorescence hilo endomicroscopy", *J Biomed Opt*, vol. 17, no. 2, p. 021 105, 2012, ISSN: 1560-2281 (Electronic) 1083-3668 (Linking). DOI: 10.1117/1.JBO.17.2.021105. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/22463023.

[87] K. Vyas, M. Hughes, B. G. Rosa, and G. Z. Yang, "Fiber bundle shifting endomicroscopy for high-resolution imaging", *Biomed Opt Express*, vol. 9, no. 10, pp. 4649–4664, 2018, ISSN: 2156-7085 (Print) 2156-7085 (Linking). DOI: 10.1364/BOE.9.004649. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/30319893.

[88] S. C. Park, M. K. Park, and M. G. Kang, "Super-resolution image reconstruction: A technical overview", *IEEE signal processing magazine*, vol. 20, no. 3, pp. 21–36, 2003, ISSN: 1053-5888.

[89] A. Orth, M. Ploschner, I. S. Maksymov, and B. C. Gibson, "Extended depth of field imaging through multicore optical fibers", *Opt Express*, vol. 26, no. 5, pp. 6407–6419, 2018, ISSN: 1094-4087 (Electronic) 1094-4087 (Linking). DOI: 10.1364/OE.26.006407. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/29529833.

[90] A. Orth, M. Ploschner, E. R. Wilson, I. S. Makysymov, and B. C. Gibson, "Optical fiber bundles: Ultra-slim light field imaging probe", *Science Advances*, 2019.

[91] Generic, 2007.

[92] H. Neumann, F. S. Fuchs, M. Vieth, R. Atreya, J. Siebler, R. Kiesslich, and M. Neurath, "In vivo imaging by endocytoscopy", *Alimentary pharmacology and therapeutics*, vol. 33, no. 11, pp. 1183–1193, 2011, ISSN: 0269-2813.

[93] T. Ohigashi, N. Kozakai, R. Mizuno, A. Miyajima, and M. Murai, "Endocytoscopy: Novel endoscopic imaging technology for in-situ observation of bladder cancer cells", *Journal of Endourology*, vol. 20, no. 9, pp. 698–701,

2006, ISSN: 0892-7790. DOI: 10.1089/end.2006.20.698. [Online]. Available: https://doi.org/10.1089/end.2006.20.698.

[94] M. Hughes, P. Giataganas, and G. Z. Yang, "Color reflectance fiber bundle endomicroscopy without back-reflections", *J Biomed Opt*, vol. 19, no. 3, p. 30 501, 2014, ISSN: 1560-2281 (Electronic) 1083-3668 (Linking). DOI: 10.1117/1.JBO.19.3.030501. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/24589984.

[95] G. J. Tearney, R. H. Webb, and B. E. Bouma, "Spectrally encoded confocal microscopy", *Opt Lett*, vol. 23, no. 15, pp. 1152–4, 1998, ISSN: 0146-9592 (Print) 0146-9592 (Linking). DOI: 10.1364/ol.23.001152. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/18087457.

[96] D. Yelin, W. M. White, J. T. Motz, S. H. Yun, B. E. Bouma, and G. J. Tearney, "Spectral-domain spectrally-encoded endoscopy", *Optics Express*, vol. 15, no. 5, p. 2432, 2007, ISSN: 1094-4087. DOI: 10.1364/oe.15.002432. [Online]. Available: https://dx.doi.org/10.1364/oe.15.002432.

[97] M. Strupler, E. D. Montigny, D. Morneau, and C. Boudoux, "Rapid spectrally encoded fluorescence imaging using a wavelength-swept source", *Optics Letters*, vol. 35, 2010.

[98] J. T. Motz, D. Yelin, B. J. Vakoc, B. E. Bouma, and G. J. Tearney, "Spectral- and frequency-encoded fluorescence imaging", *Opt Lett*, vol. 30, no. 20, pp. 2760–2, 2005, ISSN: 0146-9592 (Print) 0146-9592 (Linking). DOI: 10.1364/ol.30.002760. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/16252766.

[99] N. Bedard and T. S. Tkaczyk, "Snapshot spectrally encoded fluorescence imaging through a fiber bundle", *J Biomed Opt*, vol. 17, no. 8, pp. 080 508–1, 2012, ISSN: 1560-2281 (Electronic) 1083-3668 (Linking). DOI: 10.1117/1.JBO.17.8.080508. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/23224159.

[100] D. Y. Kim, K. Hwang, J. Ahn, Y.-H. Seo, J.-B. Kim, S. Lee, J.-H. Yoon, E. Kong, Y. Jeong, S. Jon, P. Kim, and K.-H. Jeong, "Lissajous scanning two-photon endomicroscope for in vivo tissue imaging", *Scientific Reports*, vol. 9, no. 1, 2019, ISSN: 2045-2322. DOI: 10.1038/s41598-019-38762-w. [Online]. Available: https://dx.doi.org/10.1038/s41598-019-38762-w%20https://kp-pdf.s3.amazonaws.com/d4773f11-fd1d-40c0-8d4e-e1eb3c60f98c.pdf?X-Amz-Algorithm=AWS4-HMAC-SHA256&X-Amz-Credential=AKIAUROH2NUQSIQZIEG4%2F20200615%2Fus-east-1%2Fs3%2Faws4_request&X-Amz-

Date=20200615T144254Z&X-Amz-Expires=600&X-Amz-SignedHeaders=host&X-Amz-Signature=be9cf765ba0cea5491a03f41a6a95f4f9407e6e3b

[101] E. Zagaynova, N. Gladkova, N. Shakhova, G. Gelikonov, and V. Gelikonov, "Endoscopic oct with forward-looking probe: Clinical studies in urology and gastroenterology", *Journal of biophotonics*, vol. 1, no. 2, pp. 114–128, 2008, ISSN: 1864-063X.

[102] A. G. Podoleanu, "Optical coherence tomography", *J Microsc*, vol. 247, no. 3, pp. 209–19, 2012, ISSN: 1365-2818 (Electronic) 0022-2720 (Linking). DOI: 10.1111/j.1365-2818.2012.03619.x. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/22708800.

[103] B. A. Flusberg, J. C. Jung, E. D. Cocker, E. P. Anderson, and M. J. Schnitzer, "In vivo brain imaging using a portable 3.9?gram two-photon fluorescence microendoscope", *Optics Letters*, vol. 30, no. 17, pp. 2272–2274, 2005. DOI: 10.1364/OL.30.002272. [Online]. Available: http://ol.osa.org/abstract.cfm?URI=ol-30-17-2272.

[104] W. Piyawattanametha, E. D. Cocker, L. D. Burns, R. P. J. Barretto, J. C. Jung, H. Ra, O. Solgaard, and M. J. Schnitzer, "In vivo brain imaging using a portable 29 g two-photon microscope based on a microelectromechanical systems scanning mirror", *Optics Letters*, vol. 34, no. 15, p. 2309, 2009, ISSN: 0146-9592. DOI: 10.1364/ol.34.002309. [Online]. Available: https://dx.doi.org/10.1364/OL.34.002309%20https://kp-pdf.s3.amazonaws.com/96bbb2e7-2a44-485f-b4ff-78d8e31e5cf7.pdf?X-Amz-Algorithm=AWS4-HMAC-SHA256&X-Amz-Credential=AKIAUROH2NUQSIQZIEG4%2F20200615%2Fus-east-1%2Fs3%2Faws4_request&X-Amz-Date=20200615T144728Z&X-Amz-Expires=600&X-Amz-SignedHeaders=host&X-Amz-Signature=59c9ea1ea3d2ef1ef3645d7e79338364a6aafa329e804e82e405

[105] M. T. Myaing, D. J. MacDonald, and X. Li, "Fiber-optic scanning two-photon fluorescence endoscope", *Optics letters*, vol. 31, no. 8, pp. 1076–1078, 2006, ISSN: 1539-4794.

[106] W. Piyawattanametha, R. P. J. Barretto, T. H. Ko, B. A. Flusberg, E. D. Cocker, H. Ra, D. Lee, O. Solgaard, and M. J. Schnitzer, "Fast-scanning two-photon fluorescence imaging based on a microelectromechanical systems two- dimensional scanning mirror", *Optics letters*, vol. 31, no. 13, pp. 2018–2020, 2006, ISSN: 0146-9592 1539-4794. DOI: 10.1364/ol.31.002018. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/

16770418%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2820363/.

[107] W. Denk, J. Strickler, and W. Webb, "Two-photon laser scanning fluorescence microscopy", *Science*, vol. 248, no. 4951, pp. 73–76, 1990, ISSN: 0036-8075. DOI: 10.1126/science.2321027. [Online]. Available: https://dx.doi.org/10.1126/science.2321027.

[108] H. Mehidine, M. Li, J.-F. Lendresse, F. Bouvet, H. Xie, and D. Abi Haidar, "A customized two photon fluorescence imaging probe based on 2d scanning mems mirror including electrothermal two-level-ladder dual s-shaped actuators", *Micromachines*, vol. 11, no. 7, p. 704, 2020.

[109] C. Wang and N. Ji, "Characterization and improvement of three-dimensional imaging performance of grin-lens-based two-photon fluorescence endomicroscopes with adaptive optics", *Optics Express*, vol. 21, no. 22, pp. 27 142–27 154, 2013, ISSN: 1094-4087.

[110] W. Liang, H.-C. Park, K. Li, A. Li, D. Chen, H. Guan, Y. Yue, Y.-T. A. Gau, D. E. Bergles, and M.-J. Li, "Throughput-speed product augmentation for scanning fiber-optic two-photon endomicroscopy", *IEEE Transactions on Medical Imaging*, 2020, ISSN: 0278-0062.

[111] H. Mehidine, M. Li, J.-F. Lendresse, F. Bouvet, H. Xie, and D. Abi Haidar, "A customized two photon fluorescence imaging probe based on 2d scanning mems mirror including electrothermal two-level-ladder dual s-shaped actuators", *Micromachines*, vol. 11, no. 7, p. 704, 2020.

[112] G. Li, X. Duan, M. Lee, M. Birla, J. Chen, K. R. Oldham, T. D. Wang, and H. Li, "Ultra-compact microsystems-based confocal endomicroscope", *IEEE Transactions on Medical Imaging*, 2020, ISSN: 0278-0062.

[113] G. Matz, B. Messerschmidt, W. Göbel, S. Filser, C. S. Betz, M. Kirsch, O. Uckermann, M. Kunze, S. Flämig, and A. Ehrhardt, "Chip-on-the-tip compact flexible endoscopic epifluorescence video-microscope for in-vivo imaging in medicine and biomedical research", *Biomedical Optics Express*, vol. 8, no. 7, pp. 3329–3342, 2017, ISSN: 2156-7085.

[114] D. C. Harris, "A peek into the history of sapphire crystal growth", in *Window and Dome Technologies VIII*, vol. 5078, International Society for Optics and Photonics, pp. 1–11.

[115] A. Orth and K. B. Crozier, "Light field moment imaging", *Opt Lett*, vol. 38, no. 15, pp. 2666–8, 2013, ISSN: 1539-4794 (Electronic) 0146-9592 (Linking). DOI: 10.1364/OL.38.002666. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/23903106.

[116] T. N. Ford, K. K. Chu, and J. Mertz, "Phase-gradient microscopy in thick tissue with oblique back-illumination", *Nature Methods*, vol. 9, no. 12, pp. 1195–1197, 2012, ISSN: 1548-7091 1548-7105. DOI: 10.1038/nmeth.2219.

[117] H. P. Paudel, C. Alt, J. Runnels, and C. P. Lin, "Pupil plane differential detection microscopy", *Opt Lett*, vol. 43, no. 18, pp. 4410–4412, 2018, ISSN: 1539-4794 (Electronic) 0146-9592 (Linking). DOI: 10.1364/OL.43.004410. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/30211877.

[118] J. H. Han, J. Lee, and J. U. Kang, "Pixelation effect removal from fiber bundle probe based optical coherence tomography imaging", *Opt Express*, vol. 18, no. 7, pp. 7427–39, 2010, ISSN: 1094-4087 (Electronic) 1094-4087 (Linking). DOI: 10.1364/OE.18.007427. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/20389766.

[119] S. Qu, H. Zhou, R. Liu, Y. Chen, S. Zu, S. Yu, J. Yuan, and Y. Yang, "Deblending of simultaneous-source seismic data using fast iterative shrinkage-thresholding algorithm with firm-thresholding", *Acta Geophysica*, vol. 64, no. 4, pp. 1064–1092, 2016, ISSN: 1895-6572 1895-7455. DOI: 10.1515/acgeo-2016-0043.

[120] K. E. Loewke, D. B. Camarillo, C. A. Jobst, and J. K. Salisbury, "Real-time image mosaicing for medical applications", *Stud Health Technol Inform*, vol. 125, pp. 304–9, 2007, ISSN: 0926-9630 (Print) 0926-9630 (Linking). [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/17377290.

[121] K. E. Loewke, D. B. Camarillo, W. Piyawattanametha, M. J. Mandella, C. H. Contag, S. Thrun, and J. K. Salisbury, "In vivo micro-image mosaicing", *IEEE Trans Biomed Eng*, vol. 58, no. 1, pp. 159–71, 2011, ISSN: 1558-2531 (Electronic) 0018-9294 (Linking). DOI: 10.1109/TBME.2010.2085082. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/20934939.

[122] M. Kyrish, R. Kester, R. Richards-Kortum, and T. Tkaczyk, "Improving spatial resolution of a fiber bundle optical biopsy system", *Proc SPIE Int Soc Opt Eng*, vol. 7558, p. 755 807, 2010, ISSN: 0277-786X (Print) 0277-786X

(Linking). DOI: 10.1117/12.842744. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/20711486.

[123] A. Perperidis, K. Dhaliwal, S. McLaughlin, and T. Vercauteren, "Image computing for fibre-bundle endomicroscopy: A review", 2018.

[124] P. Aslan, R. L. Kuo, K. Hazel, R. K. Babayan, and G. M. Preminger, "Advances in digital imaging during endoscopic surgery", *Journal of Endourology*, vol. 13, no. 4, pp. 251–255, 1999, ISSN: 0892-7790. DOI: 10.1089/end.1999.13.251. [Online]. Available: https://doi.org/10.1089/end.1999.13.251.

[125] J. Liu, W. Zhou, B. Xu, X. Yang, and D. Xiong, "Honeycomb pattern removal for fiber bundle endomicroscopy based on a two-step iterative shrinkage thresholding algorithm", *AIP Advances*, vol. 10, no. 4, p. 045 004, 2020, ISSN: 2158-3226.

[126] M. M. Dickens, D. J. Bornhop, and S. Mitra, "Removal of optical fiber interference in color micro-endoscopic images", IEEE Comput. Soc. DOI: 10.1109/cbms.1998.701364. [Online]. Available: https://dx.doi.org/10.1109/CBMS.1998.701364.

[127] M. Elter, S. Rupp, and C. Winter, "Physically motivated reconstruction of fiberscopic images", IEEE. DOI: 10.1109/icpr.2006.932. [Online]. Available: https://dx.doi.org/10.1109/ICPR.2006.932.

[128] K. Mulmuley, "Computational geometry", *An Introduction Through Randomized Algorithms. Prentice-Hall*, 1994.

[129] J. P. Lewis, "Fast normalized cross-correlation",

[130] ——, "Fast template matching", *Vision Interface 95, Canadian Image Processing and Pattern Recognition Society*, 1995.

[131] 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, no. 24, pp. 1905–1907, 1997, ISSN: 0146-9592. DOI: Doi10.1364/Ol.22.001905. [Online]. Available: %3CGo%20to%20ISI%3E://WOS:000071030800026.

[132] W. Wiegand, A. A. Thaer, P. Kroll, O. C. Geyer, and A. J. Garcia, "Optical sectioning of the cornea with a new confocal in vivo slit-scanning videomicroscope", *Ophthalmology*, vol. 102, no. 4, pp. 568–575, 1995, ISSN: 01616420. DOI: 10.1016/s0161-6420(95)30981-5.

[133] N. Burkhard, "Simultaneous amplitude modulation contrast (samc) a new optical imaging technique", in *Proc.SPIE*, vol. 1139. DOI: 10.1117/12.961781. [Online]. Available: https://doi.org/10.1117/12.961781.

[134] N. Hagen, L. Gao, and T. S. Tkaczyk, "Quantitative sectioning and noise analysis for structured illumination microscopy", *Opt Express*, vol. 20, no. 1, pp. 403–13, 2012, ISSN: 1094-4087 (Electronic) 1094-4087 (Linking). DOI: 10.1364/OE.20.000403. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/22274364.

[135] D. Karadaglic and T. Wilson, "Image formation in structured illumination wide-field fluorescence microscopy", *Micron*, vol. 39, no. 7, pp. 808–18, 2008, ISSN: 0968-4328 (Print) 0968-4328 (Linking). DOI: 10.1016/j.micron.2008.01.017. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/18337108.

[136] T. Wilson, "Optical sectioning in fluorescence microscopy", *J Microsc*, vol. 242, no. 2, pp. 111–6, 2011, ISSN: 1365-2818 (Electronic) 0022-2720 (Linking). DOI: 10.1111/j.1365-2818.2010.03457.x. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/21118248.

[137] M. A. A. Neil, R. Juškaitis, and T. Wilson, "Real time 3d fluorescence microscopy by two beam interference illumination", *Optics Communications*, vol. 153, no. 1, pp. 1–4, 1998, ISSN: 0030-4018. DOI: https://doi.org/10.1016/S0030-4018(98)00210-7. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0030401898002107.

[138] L. H. SCHAEFER, D. SCHUSTER, and J. SCHAFFER, "Structured illumination microscopy: Artefact analysis and reduction utilizing a parameter optimization approach", *Journal of Microscopy*, vol. 216, no. 2, pp. 165–174, 2004, ISSN: 0022-2720. DOI: 10.1111/j.0022-2720.2004.01411.x. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.0022-2720.2004.01411.x.

[139] M. Cole, J. Siegel, S. Webb, R. Jones, K. Dowling, M. J. Dayel, D. Parsons-Karavassilis, P. French, M. Lever, and L. Sucharov, "Time-domain whole-field fluorescence lifetime imaging with optical sectioning", *Journal of Microscopy*, vol. 203, no. 3, pp. 246–257, 2001, ISSN: 0022-2720.

[140] F. Chasles, B. Dubertret, and A. C. Boccara, "Optimization and characterization of a structured illumination microscope", *Opt Express*, vol. 15,

no. 24, pp. 16 130–40, 2007, ISSN: 1094-4087 (Electronic) 1094-4087 (Linking). [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/19550902.

[141] D. Xu, T. Jiang, A. Li, B. Hu, Z. Feng, H. Gong, S. Zeng, and Q. Luo, "Fast optical sectioning obtained by structured illumination microscopy using a digital mirror device", *J Biomed Opt*, vol. 18, no. 6, p. 060 503, 2013, ISSN: 1560-2281 (Electronic) 1083-3668 (Linking). DOI: 10.1117/1.JBO.18.6.060503. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/23757041.

[142] A. G. York, S. H. Parekh, D. D. Nogare, R. S. Fischer, K. Temprine, M. Mione, A. B. Chitnis, C. A. Combs, and H. Shroff, "Resolution doubling in live, multicellular organisms via multifocal structured illumination microscopy", *Nature Methods*, vol. 9, no. 7, pp. 749–754, 2012, ISSN: 1548-7091. DOI: 10.1038/nmeth.2025. [Online]. Available: https://dx.doi.org/10.1038/nmeth.2025.

[143] T. Instruments, "Dlp® lightcrafter™ dm365 command interface guide", 2013.

[144] S. Gruppetta and S. Chetty, "Theoretical study of multispectral structured illumination for depth resolved imaging of non-stationary objects: Focus on retinal imaging", *Biomedical Optics Express*, vol. 2, no. 2, pp. 255–263, 2011, ISSN: 2156-7085.

[145] B. A. Flusberg, E. D. Cocker, W. Piyawattanametha, J. C. Jung, E. L. Cheung, and M. J. Schnitzer, "Fiber-optic fluorescence imaging", *Nat Methods*, vol. 2, no. 12, pp. 941–50, 2005, ISSN: 1548-7091 (Print) 1548-7091 (Linking). DOI: 10.1038/nmeth820. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/16299479.

[146] B. Rosa, B. Dahroug, B. Tamadazte, K. Rabenorosoa, P. Rougeot, N. Andreff, and P. Renaud, "Online robust endomicroscopy video mosaicking using robot prior", *IEEE Robotics and Automation Letters*, vol. 3, no. 4, pp. 4163–4170, 2018, ISSN: 2377-3766 2377-3774. DOI: 10.1109/lra.2018.2863372.

[147] P. Wisanuvej, P. Giataganas, K. Leibrandt, J. Liu, M. Hughes, and G. Yang, "Three-dimensional robotic-assisted endomicroscopy with a force adaptive robotic arm", in *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 2379–2384. DOI: 10.1109/ICRA.2017.7989276.

[148] S. Zuo, M. Hughes, and G.-Z. Yang, "Flexible robotic scanning device for intraoperative endomicroscopy in mis", *IEEE/ASME Transactions on Mechatronics*, vol. 22, no. 4, pp. 1728–1735, 2017, ISSN: 1083-4435.

[149] N. Bedard, T. Quang, K. Schmeler, R. Richards-Kortum, and T. S. Tkaczyk, "Real-time video mosaicing with a high-resolution microendoscope", *Biomed Opt Express*, vol. 3, no. 10, pp. 2428–35, 2012, ISSN: 2156-7085 (Electronic) 2156-7085 (Linking). DOI: 10.1364/BOE.3.002428. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/23082285.

[150] A. Shinde and M. V. Matham, "Pixelate removal in an image fiber probe endoscope incorporating comb structure removal methods", *Journal of Medical Imaging and Health Informatics*, vol. 4, no. 2, pp. 203–211, 2014, ISSN: 21567018 21567026. DOI: 10.1166/jmihi.2014.1255.

[151] B. Obstoy, M. Salaun, L. Veresezan, R. Sesboue, P. Bohn, F. X. Boland, and L. Thiberville, "Safety and performance analysis of acriflavine and methylene blue for in vivo imaging of precancerous lesions using fibered confocal fluorescence microscopy (fcfm): An experimental study", *BMC Pulm Med*, vol. 15, p. 30, 2015, ISSN: 1471-2466 (Electronic) 1471-2466 (Linking). DOI: 10.1186/s12890-015-0020-4. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/25880748.

[152] C. Ventalon and J. Mertz, "Dynamic speckle illumination microscopy with translated versus randomized speckle patterns", *Opt Express*, vol. 14, no. 16, pp. 7198–209, 2006, ISSN: 1094-4087 (Electronic) 1094-4087 (Linking). DOI: 10.1364/oe.14.007198. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/19529088.

[153] Z. R. Hoffman and C. A. DiMarzio, "Structured illumination microscopy using random intensity incoherent reflectance", *J Biomed Opt*, vol. 18, no. 6, p. 061 216, 2013, ISSN: 1560-2281 (Electronic) 1083-3668 (Linking). DOI: 10.1117/1.JBO.18.6.061216. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/23183657.

[154] A. Vigoren and J. M. Zavislan, "Optical sectioning enhancement using higher-order moment signals in random speckle-structured illumination microscopy", *Journal of the Optical Society of America a-Optics Image Science and Vision*, vol. 35, no. 3, pp. 474–479, 2018, ISSN: 1084-7529. DOI: 10.1364/josaa.35.000474. [Online]. Available: %3CGo%20to%20ISI%3E://WOS:000427443300014.

[155] C. Ventalon and J. Mertz, "Quasi-confocal fluorescence sectioning with dynamic speckle illumination", *Optics letters*, vol. 30, no. 24, pp. 3350–3352, 2005, ISSN: 1539-4794.

[156] N. Bender, H. Yılmaz, Y. Bromberg, and H. Cao, "Creating and controlling complex light", *APL Photonics*, vol. 4, no. 11, p. 110 806, 2019, ISSN: 2378-0967. DOI: 10.1063/1.5132960. [Online]. Available: https://dx.doi.org/10.1063/1.5132960.

[157] F. Meng, Y. Zhao, Y. Z. Zhang, and L. Huo, "Tuning the intensity statistics of random speckle patterns", *Chinese Physics B*, vol. 28, no. 5, p. 5, 2019, ISSN: 1674-1056. DOI: 10.1088/1674-1056/28/5/057801. [Online]. Available: %3CGo%20to%20ISI%3E://WOS:000468333100001.

[158] N. Bender, H. Yilmaz, Y. Bromberg, and H. Cao, "Customizing speckle intensity statistics", *Optica*, vol. 5, no. 5, pp. 595–600, 2018, ISSN: 2334-2536. DOI: 10.1364/optica.5.000595. [Online]. Available: %3CGo%20to%20ISI%3E://WOS:000432560400013.

[159] C. J. R. Sheppard and X. Q. Mao, "Confocal microscopes with slit apertures", *Journal of Modern Optics*, vol. 35, no. 7, pp. 1169–1185, 1988, ISSN: 0950-0340 1362-3044. DOI: 10.1080/09500348814551251.

[160] P. J. Timoney and C. S. Breathnach, "Allvar gullstrand and the slit lamp 1911", *Irish Journal of Medical Science*, vol. 182, no. 2, pp. 301–305, 2013, ISSN: 0021-1265. DOI: 10.1007/s11845-012-0873-y. [Online]. Available: https://dx.doi.org/10.1007/s11845-012-0873-y%20https://kp-pdf.s3.amazonaws.com/0989a9e0-fd46-4fb4-9b40-b55282c62ad3.pdf?X-Amz-Algorithm=AWS4-HMAC-SHA256&X-Amz-Credential=AKIAUROH2NUQSIQZIEG4%2F20200610%2Fus-east-1%2Fs3%2Faws4_request&X-Amz-Date=20200610T113608Z&X-Amz-Expires=600&X-Amz-SignedHeaders=host&X-Amz-Signature=5aa7cce5ab6a7b6f052796dc05b8ce912bdeffeb476b25e5c13c

[161] E. F. Fincham, "A new form of corneal microscope with combined slit lamp illuminating device", *Transactions of the Optical Society*, vol. 25, no. 3, pp. 113–122, 1924, ISSN: 1475-4878. DOI: 10.1088/1475-4878/25/3/303. [Online]. Available: https://dx.doi.org/10.1088/1475-4878/25/3/303%20https://kp-pdf.s3.amazonaws.com/24ce490e-439b-4fab-b53f-c97a70458780.pdf?X-Amz-Algorithm=AWS4-HMAC-SHA256&X-Amz-Credential=AKIAUROH2NUQSIQZIEG4%2F20200610%2Fus-east-1%2Fs3%2Faws4_request&X-Amz-

Date=20200610T121709Z&X-Amz-Expires=600&X-Amz-SignedHeaders=host&X-Amz-Signature=ed7594b32dea09e7d1a9766607d9f24a579460f4c

[162] D. M. Maurice, "A scanning slit optical microscope", *Investigative ophthalmology*, vol. 13, no. 12, pp. 1033–1037, 1974, ISSN: 0020-9988. [Online]. Available: http://europepmc.org/abstract/MED/4430572.

[163] C. J. Koester, "Scanning mirror microscope with optical sectioning characteristics: Applications in ophthalmology", vol. 19, no. 11, p. 1749, 1980, ISSN: 0003-6935. DOI: 10.1364/ao.19.001749. [Online]. Available: https://dx.doi.org/10.1364/AO.19.001749%20https://kp-pdf.s3.amazonaws.com/2b1be508-a74b-4847-bed0-f2974f3183b0.pdf?X-Amz-Algorithm=AWS4-HMAC-SHA256&X-Amz-Credential=AKIAUROH2NUQSIQZIEG4%2F20200609%2Fus-east-1%2Fs3%2Faws4_request&X-Amz-Date=20200609T164403Z&X-Amz-Expires=600&X-Amz-SignedHeaders=host&X-Amz-Signature=da2cfb41efbc4085349edb80005996ea5e226bb0e01ca8821c73

[164] T. Wilson and A. R. Carlini, "Three-dimensional imaging in confocal imaging systems with finite sized detectors", *Journal of Microscopy*, vol. 149, no. 1, pp. 51–66, 1988, ISSN: 0022-2720. DOI: https://doi.org/10.1111/j.1365-2818.1988.tb04561.x. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2818.1988.tb04561.x.

[165] R. Gauderon and C. J. R. Sheppard, "Improvement in imaging in confocal fluorescent microscopes using multiple detectors", *Bioimaging*, vol. 6, no. 3, pp. 126–129, 1998, ISSN: 0966-9051. DOI: 10.1002/1361-6374(199809)6:3<126::aid-bio2>3.0.co;2-l. [Online]. Available: https://dx.doi.org/10.1002/1361-6374(199809)6:3%3C126::aid-bio2%3E3.0.co;2-l%20https://kp-pdf.s3.amazonaws.com/782de6e0-cc8c-4e80-b448-f54584d7493e.pdf?X-Amz-Algorithm=AWS4-HMAC-SHA256&X-Amz-Credential=AKIAUROH2NUQSIQZIE2F20200612%2Fus-east-1%2Fs3%2Faws4_request&X-Amz-Date=20200612T115505Z&X-Amz-Expires=600&X-Amz-SignedHeaders=host&X-Amz-Signature=e680cbd8ae95d3e130f5b614f0c7f8213890a695d

[166] D. Guang, "A generalized unsharp masking algorithm", *IEEE Transactions on Image Processing*, vol. 20, no. 5, pp. 1249–1261, 2011, ISSN: 1057-7149. DOI: 10.1109/tip.2010.2092441. [Online]. Available: https://dx.doi.org/10.1109/tip.2010.2092441.

[167] B. R. Masters and A. A. Thaer, "Real-time scanning slit confocal microscopy of the in vivo human cornea", *Appl Opt*, vol. 33, no. 4, pp. 695–701, 1994, ISSN: 1559-128X (Print) 1559-128X (Linking). DOI: 10.1364/AO.33.000695. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/20862066.

[168] Y. S. Sabharwal, A. R. Rouse, L. Donaldson, M. F. Hopkins, and A. F. Gmitro, "Slit-scanning confocal microendoscope for high-resolution in vivo imaging", *Appl Opt*, vol. 38, no. 34, pp. 7133–44, 1999, ISSN: 1559-128X (Print) 1559-128X (Linking). DOI: 10.1364/ao.38.007133. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/18324260.

[169] E. Mei, P. Fomitchov, R. Graves, and M. Campion, "A line scanning confocal fluorescent microscope using a cmos rolling shutter as an adjustable aperture", *Journal of microscopy*, vol. 247, no. 3, pp. 269–276, 2012, ISSN: 0022-2720.