Multi-modal Whole Slide Imaging

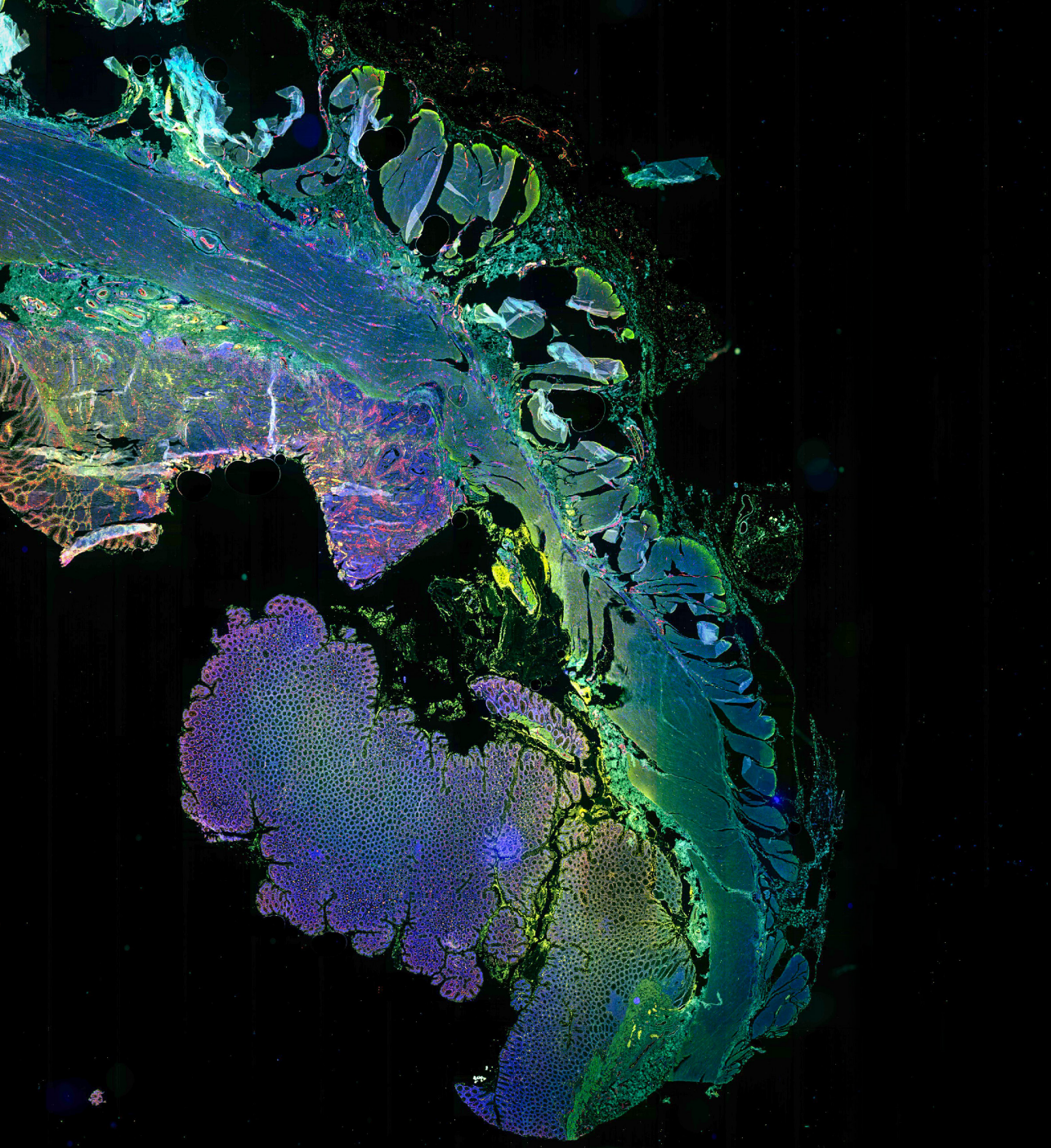
van der Graaff, Leon

DOI
10.4233/uuid:056d0084-880b-438f-9589-e41a108b425a

Publication date
2020

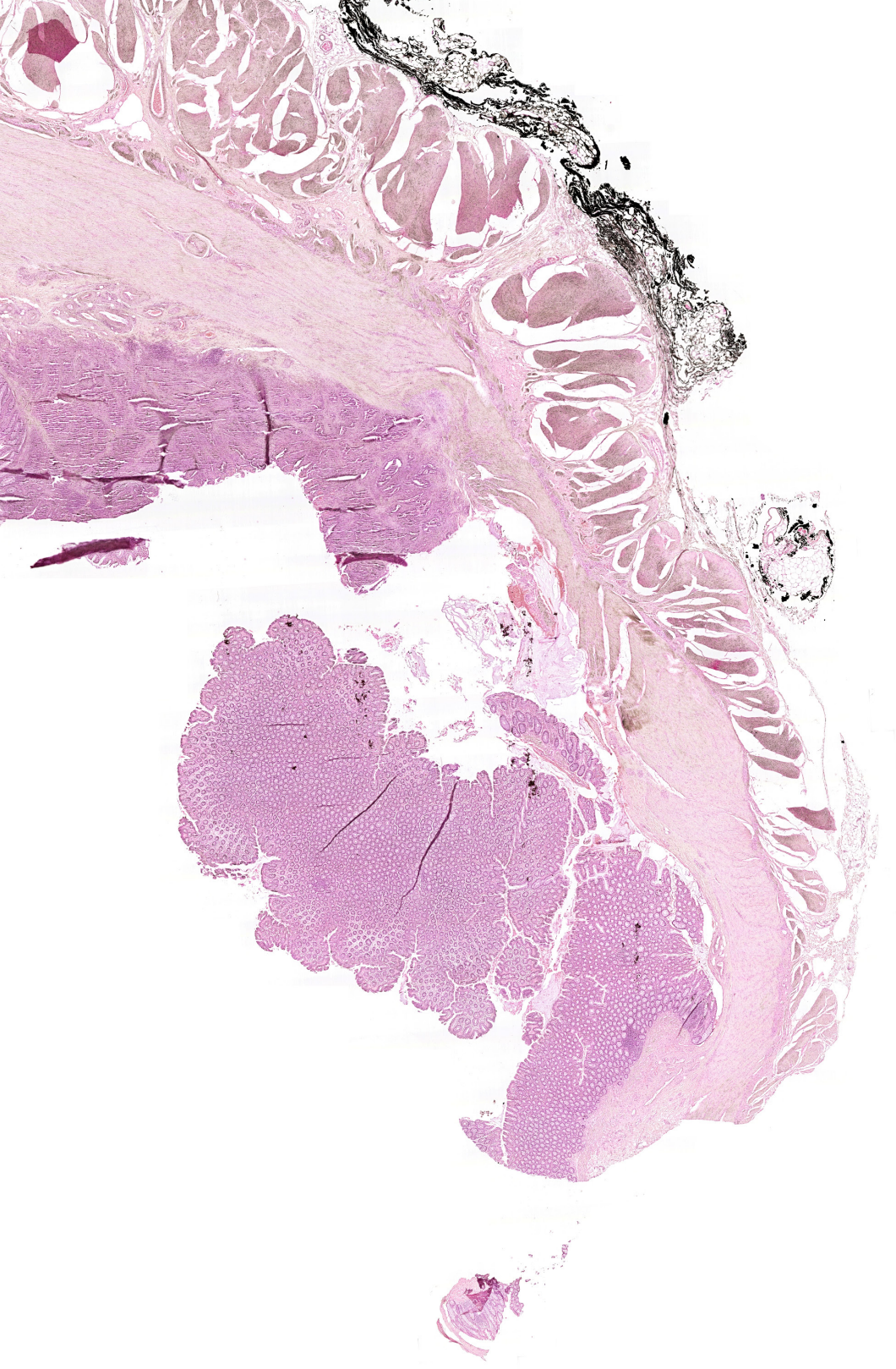
Document Version
Final published version

Citation (APA)
van der Graaff, L. (2020). *Multi-modal Whole Slide Imaging*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:056d0084-880b-438f-9589-e41a108b425a



Multi-modal Whole Slide Imaging

Leon van der Graaff



TU Delft ImPhys

ISBN 978-94-6384-158-0

Multi-modal Whole Slide Imaging

Multi-modal Whole Slide Imaging

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology
by the authority of the Rector Magnificus, prof. dr. ir. T.H.J.J. van der Hagen,
chair of the Board for Doctorates
to be defended publicly on
Wednesday 9, September 2020 at 15:00 o'clock

by

Leon VAN DER GRAAFF

Master of Science in Applied Physics,
Delft University of Technology, Delft, The Netherlands,
born in Rotterdam, The Netherlands.

This dissertation has been approved by the promotors.

Composition of the doctoral committee:

Rector Magnificus,	chairperson
Prof. dr. S. Stallinga,	Delft University of Technology, promotor
Prof. dr. ir. L.J. van Vliet,	Delft University of Technology, promotor
Independent members:	
Prof. dr. M.L. Groot,	Vrije Universiteit Amsterdam
Prof. dr. T. Lasser,	École Polytechnique Fédérale de Lausanne, Switzerland
Prof. dr. H.P. Urbach,	Delft University of Technology
Dr. ir. P. Zijlstra,	Eindhoven University of Technology
Prof. dr. B. Rieger,	Delft University of Technology, reserve member
Other members:	
Dr. G.J.L.H. van Leenders,	Erasmus University Medical Center
Dr. ir. J.P. Vink,	Philips Digital & Computational Pathology

The work in this thesis was conducted at the Imaging Physics department, faculty of Applied Sciences, Delft University of Technology.
This research was publicly funded by NWO-TTW (Nederlandse Organisatie voor Wetenschappelijk Onderzoek - Technische en Toegepaste Wetenschappen, project 13892) with support of Philips Digital & Computational Pathology.

The cover shows two Whole Slide Images of the same tissue, acquired using different imaging modalities. Front: a multi-color widefield immunofluorescence image. Back: H&E brightfield image. Two neighbouring 4 µm thick stage 3 human rectum cancer sections are used. The samples are kindly provided by Fanny Boyaval. Credits to Philips for the technical support.

Printed by: Ipskamp Printing

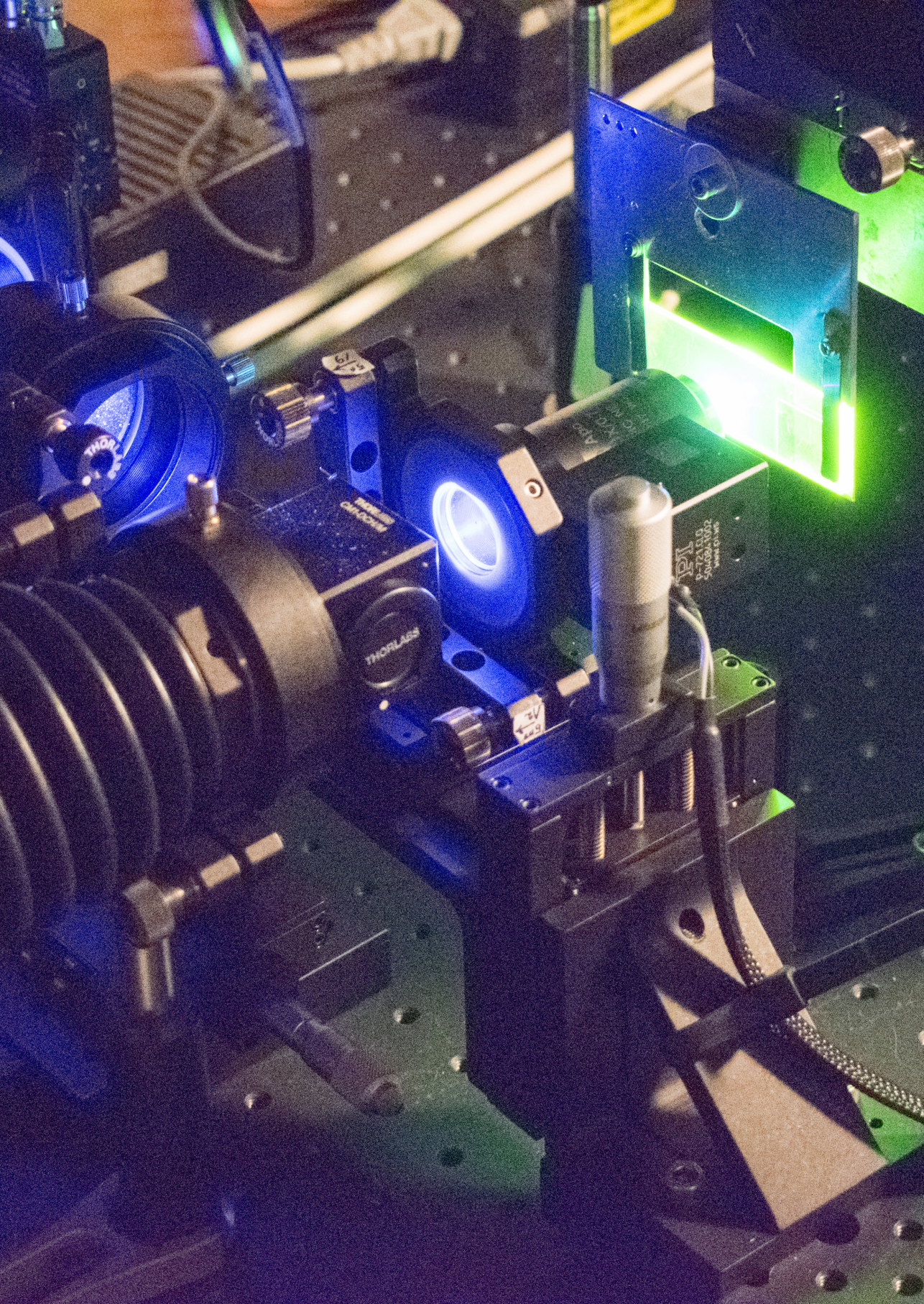
978-94-6384-158-0

An electronic version of this dissertation is available at repository.tudelft.nl.

Nu kijken we nog in een wazige spiegel, maar straks staan we oog in oog.

1 Korintiërs 13

Contents


THORLABS

Summary

This thesis explores a novel optical architecture for Whole Slide Imaging (WSI). This new architecture allows for multi-focal (3D) image acquisitions in a single scan pass. The multi-focal imaging capability is used to demonstrate 3D phase imaging and 3D imaging of thick tissue sections on a prototype scanner. Further, instrumentation for the extension of WSI to fluorescence imaging is developed: a technologically robust and cost effective method based on LEDs and a highly efficient method based on a multi-line laser illumination source.

A WSI system is an optical instrument aimed at creating digital images of biological samples mounted on a microscopy slide at a high throughput. WSI systems image tissues over large fields of view (~ few cm), in 2D or in 3D (up to hundred layers of μm thickness), and at cellular resolution (~1 μm). They are applied in high throughput screening in biology, and for novel computer aided medical diagnoses in the field of digital pathology.

The optical architecture explored in this thesis is based on a tilted multi-line image sensor concept, originating from Philips. The goal of multi-line image acquisition is to enable closed-loop autofocus scanning, but it also allows for multi-focal image acquisitions. The core of this scanner concept lies in a novel design for a multi-line image sensor. The sensor is experimentally characterized for gain and noise, and a system model is developed to find the optimum signal-to-noise ratio (SNR) given the available photo-electron flux. This showed that images with a very high SNR of 292 can be acquired, provided that a sufficiently high photon flux can be realized. Two major issues with the sensor were found in our experimental characterization. At high line rates, the sensor showed missing symbols, leading to non-linearities in the read out. Further, the sensor showed high frequent fluctuations in the gain.

3D phase imaging and 3D imaging of thick specimens are two novel contrast modalities based on computational imaging techniques and are enabled by the availability of multi-focal images. For both techniques simplified algorithms were developed compatible with parallel processing at very high speeds. For 3D imaging of thick specimens a deconvolution technique is developed for improving the inherently low axial contrast. 3D phase imaging is realized by a simplified algorithm for Quantitative Phase Tomography (QPT). QPT imaging is found to be able to image the sites labeled for Fluorescence *in situ* Hybridization (FISH) imaging, and provide additional structural information on unlabeled tissues or tissues stained for immunofluorescence. A system design study is presented showing that the in-plane transfer function has the character of a band-pass spatial frequency filter.

The major opportunity for WSI systems to become compatible with fluorescence imaging is addressed by the development of two imaging modalities. First, a widefield fluorescence WSI system with an LED illumination source is developed and built. A color sequential illumination strategy in combination with multi-band dichroics is used for multi-color imaging using a single monochromatic sensor.

The main speed limitation is formed by the exposure time required to capture enough photo-electrons for a decent SNR. Based on the experimental results, a system with 96 Time Delayed Integration (TDI) lines is estimated to achieve a reasonable throughput of about 130 kPixel/s. This makes scanning possible of an area of $15 \times 15\,\mathrm{mm}^2$ in three colors in about 23 min.

Second, a novel optical architecture for multi-focal fluorescence image acquisitions based on a laser illumination source is proposed and realized in a prototype. Illumination PSF engineering using diffractive optics is applied to generate a set of parallel scan lines in object space, that span a plane conjugate to a tilted image sensor. An important new element in the design is the use of higher order astigmatism to improve the uniformity of peak intensity and line width along the scan lines. Focusing the illumination on the sample provides a very high illumination efficiency and a confocal suppression of background. This optical architecture is projected to ultimately achieve a throughput of several hundreds MPixel/s, which would enable scanning an area of $15 \times 15\,\mathrm{mm}^2$ in 8 layers in less than a minute.

This thesis is concluded with an outlook to opportunities for future research in WSI systems. The challenges and some potential solutions for using a general purpose scientific CMOS (sCMOS) camera for multi-line scanning of a tilted object plane, and some opportunities for extension of WSI techniques to Light Sheet Microscopy (LSM) and Structured Illumination Microscopy (SIM) are discussed.

In summary, this thesis investigates the imaging qualities and extension to computational imaging modalities of a brightfield WSI system and describes two approaches for fluorescence WSI.

Samenvatting

Dit proefschrift verkent de mogelijkheden van een nieuwe optische architectuur voor 'Whole Slide Imaging' (WSI). Deze nieuwe optische architectuur maakt het tegelijkertijd opnemen van afbeeldingen op meerdere focusniveaus mogelijk. De mogelijkheid tot multi-focaal afbeelden is benut voor het demonstreren van het 3D afbeelden van fase contrast en het 3D afbeelden van dikke coupes op een prototype scanner. Verder is instrumentatie ontwikkeld voor de uitbreiding van WSI naar fluorescentie microscopie: een technologisch robuuste en kostenefficiënte methode gebaseerd op LEDs en een zeer efficiente methode gebaseerd op een multi-lijn laser belichting.

Een 'Whole Slide Scanning' WSI systeem is een optisch instrument voor het met hoge snelheid maken van digitale afbeeldingen van biologische preparaten die zijn bevestigd op een objectglaasje. WSI systemen beelden weefsels af met een groot gezichtsveld (~ een paar cm), in 2D of 3D (tot wel honderd lagen van µm dikte), en met cellulaire resolutie (~ µm). Ze worden toegepast voor 'high throughput screening' in biologie, en voor nieuwe computerondersteunde gezondheidsdiagnostiek in het vakgebied van de Digitale Pathologie.

De optische architectuur die in dit proefschrift wordt verkend is gebaseerd op een gekantelde multi-lijn beeldsensor, een concept afkomstig van Philips. Het doel van multi-lijn beeldopnamen met deze gekantelde sensor is om een gesloten feedbacksysteem voor autofocus mogelijk te maken. Daarnaast maakt het ook het tegelijkertijd opnemen van afbeeldingen op meerdere focusniveaus mogelijk. Het hart van dit concept is een nieuw ontwerp voor een multi-lijn beeldsensor. De versterking en het ruisgedrag van de sensor zijn experimenteel gekarakteriseerd, en een systeem model is ontwikkeld om de optimale signaal-ruisverhouding (SNR) te bepalen gegeven de beschikbare foto-elektronenstroom. Hieruit blijkt dat afbeeldingen met een heel hoge SNR van 292 kunnen worden verkregen als een afdoende grote foto-elektronenstroom beschikbaar is. Tijdens de experimentele karakterisatie werden twee problemen met de sensor ontdekt. Bij hoge lijnopnamesnelheden ontbreken een reeks getallen in de digitale data, wat leidt tot niet-lineariteiten. Verder werd een hoogfrequente ruisbron in de versterkingsfactor ontdekt.

Het 3D afbeelden van fase contrast en het 3D afbeelden van dikke weefselcoupes zijn twee nieuwe contrast modaliteiten gebaseerd op rekenkundige afbeeldingstechnieken die mogelijk worden gemaakt door de beschikbaarheid van multi-focale afbeeldingen. Voor beide technieken zijn vereenvoudigde algoritmes ontwikkeld die geschikt zijn voor parallelle verwerking met zeer hoge doorvoersnelheden. Voor het 3D afbeelden van dikke weefselcoupes is een deconvolutie techniek ontwikkeld ter verbetering van het inherent lage axiale contrast. Het 3D fase contrast afbeelden is gerealiseerd met een vereenvoudigd algoritme voor 'Quantitative Phase Tomography' (QPT). Ontdekt werd dat met behulp van QPT materiaal kan worden onderscheiden dat was gelabeld met 'Fluorescence *in situ*

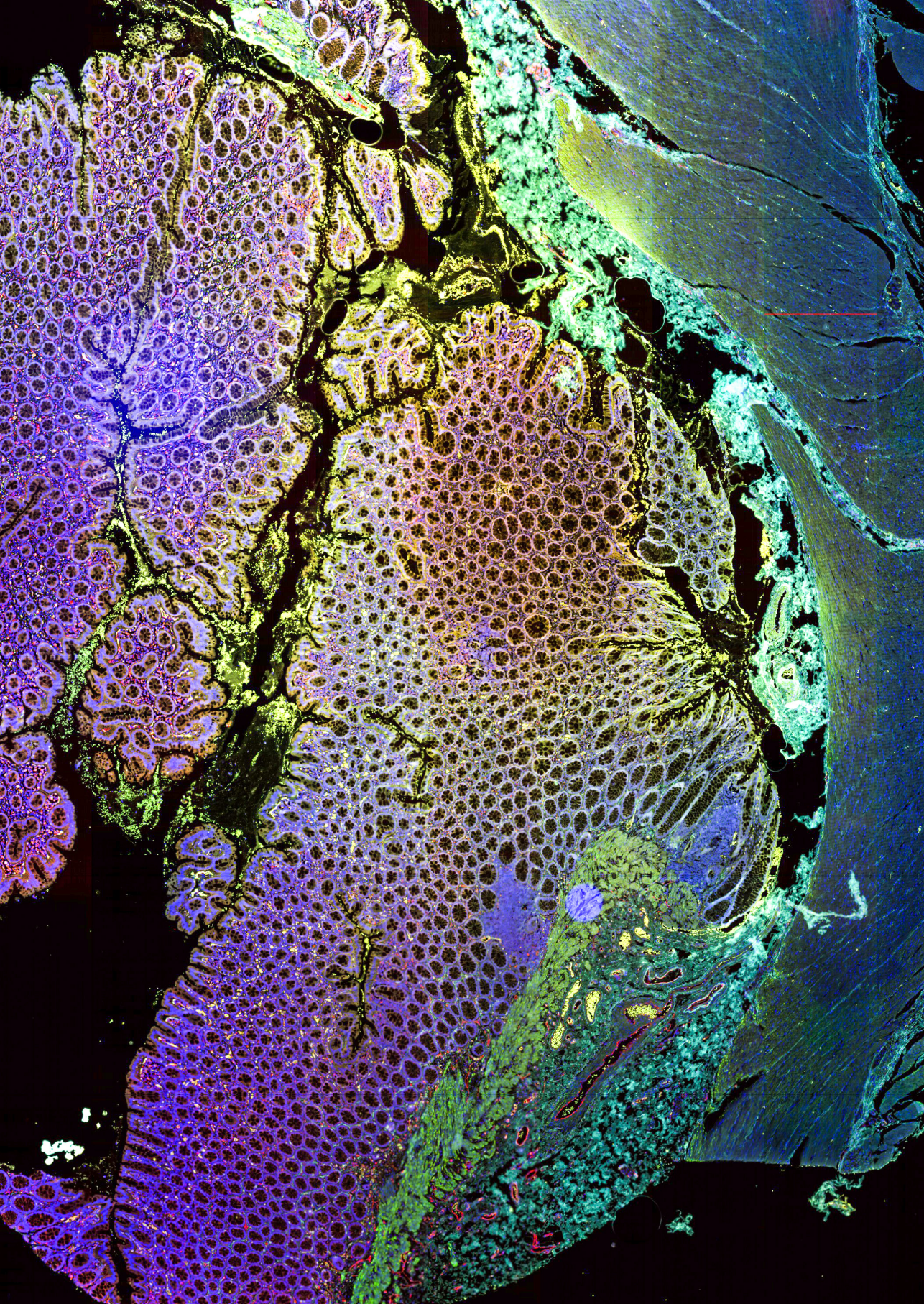
Hybridization' (FISH), en dat aanvullende structurele informatie kan worden verkregen van ongelabelde weefsels en weefsels die zijn gelabeld voor immunofluorescentie. Dit proefschrift bevat een ontwerpstudie die als conclusie heeft dat de overdrachtsfunctie in het laterale vlak het karakter heeft van een spatieel banddoorlaatfilter.

De belangrijke kans voor WSI systemen om ook fluorescentie-afbeeldingen te kunnen maken is aangegrepen om twee beeldmodaliteiten te ontwikkelen. Allereerst is een breedveld fluorescentie WSI systeem met een LED verlichting ontwikkeld en gebouwd. Een kleursequentieel belichtingsschema gecombineerd met multi-band dichroïsche spiegels zorgt ervoor dat meerdere kleurkanalen kunnen worden opgenomen met een enkele monochromatische sensor. De belangrijkste snelheidsbegrenzing is de belichtingstijd die nodig is om genoeg foto-elektronen te verzamelen voor een behoorlijke SNR. Gebaseerd op de meetresultaten wordt geschat dat een system met 96 *Time Delayed Integration* (TDI) lijnen een doorvoersnelheid van 130 kPixel/s kan halen. Dit maakt het mogelijk om een gebied van $15 \times 15\,\mathrm{mm}^2$ in drie kleuren op te nemen in ongeveer 23 min.

Ten tweede is een nieuwe architectuur voor multi-focale fluorescentie-microscopie gebaseerd op laserbelichting voorgesteld en gerealiseerd in een prototype. Technieken voor het vormgeven van de puntspreidingsfunctie van de belichting gebaseerd op diffractieve optiek werden gebruikt voor het maken van een reeks parallelle scanlijnen in de objectruimte die een vlak opspannen dat geconjugeerd is aan de gekantelde beeldsensor. Een belangrijk nieuw element in het ontwerp is het gebruik van hogere orde astigmatisme voor een verbeterde uniformiteit van de piekintensiteit en de lijnbreedte van de scanlijnen. Het focusseren van de belichting op het object geeft een zeer hoge belichtingsefficiëntie en een confocale onderdrukking van het achtergrondsignaal. Verwacht wordt dat deze optische architectuur uiteindelijk een doorvoersnelheid kan bereiken van meerdere honderden MPixel/s, wat het mogelijk maakt om een gebied van $15 \times 15\,\mathrm{mm}^2$ in 8 lagen op te nemen in minder dan een minuut.

Dit proefschrift wordt afgesloten met een vooruitblik naar nieuwe mogelijkheden van onderzoek naar WSI systemen. In het bijzonder worden de belemmeringen en enkele mogelijke oplossingen voor het gebruik van een wetenschappelijke CMOS (sCMOS) camera voor multi-lijn scannen met een gekanteld beeldvlak, en de mogelijkheden voor het gebruik van WSI technieken in 'Light Sheet Microscopy' (LSM) en 'Structured Illumination Microscopy' (SIM) besproken.

Samenvattend onderzoekt dit proefschrift de beeldkwaliteit en uitbreidingsmogelijkheden naar rekenkundige afbeeldingstechnieken van een witlicht WSI systeem en beschrijft het twee mogelijkheden voor fluorescentie-microscopie op een WSI systeem.



1

Introduction
Microscopy in the digital era

The optical microscope is a scientific instrument by nature. Its early development in the 17th century is marked by the discovery of microbiological structures by the British scientist Hooke such as the flea eye and the pores of cork, coining the word 'cell' [1, 2], and the first observation of bacteria and unicellular organisms by the Dutch tradesman Van Leeuwenhoek [2, 3]. Up to the current day, it remains one of the core instruments for biologists and pathologists for studying life and disease at the cellular level [4, 5]. The microscope is also in the public mind, probably together with safety goggles, white coats and test tubes, the most stereotypical scientific equipment [6].

1.1. Computational imaging

The application and functionality of the microscope is still a field of active research, even though it has been used in science for more than four centuries. In particular, the transition from a mere optical instrument to a digital microscope has opened a wide range of possibilities. Digital microscopy is not so much about replacing the microscope eyepiece by a camera. Rather, the camera, digital image processing and instrument control become an integral part of the imaging instrument. The field of science that uses this integrated view on instrumentation and image processing is called computational imaging, and has provided many applications in the field of microscopy, of which a few will be mentioned here as an example.

An important breakthrough made possible by the computational imaging approach, is the ability to provide a resolution exceeding the native resolution of the optical instrument. Structured Illumination Microscopy (SIM) [7] uses a series of image acquisitions with spatially periodic illumination patterns that are translated and rotated with respect to the sample, in combination with with computational reconstruction techniques to improve the resolution of the system by up to a factor two. Super-resolution optical fluctuation imaging (SOFI) [8] also provides an improved resolution. The method relies on the exploitation of the statistical properties of blinking fluorophores and does not require adjustments to the illumination. Stimulated-emission-depletion (STED) fluorescence microscopy [9] makes use of two illumination beams, one for excitation and one to deplete the area around the point of interest from fluorophores in the excited state, such that a very narrow region remains from which fluorescence signal is acquired. A scanning approach is used to create an image. In Single Molecule Localization Microscopy (SMLM) [10], the whole concept of imaging is replaced by the idea that the microscope is a measurement tool for position, intensity, and wavelength. Provided that the distribution of the light sources is sparse enough, the localization precision is no longer limited by the resolution of the instrument but rather by the number of captured photons. Digital images are then synthesized from the measured coordinates, photon counts and wavelengths, and can have a resolution that exceeds the optical instrument up to typically ten times. Sparse labeling of the sample is for example realized by using photoactivatable fluorophores in Photoactivated Localization Microscopy (PALM) [11] or by using photoswitchable fluorophores in Stochastic Optical Reconstruction Microscopy (STORM) [12]. In recognition of these breakthroughs, the Nobel prize for chemistry was awarded in 2014 for the development of super-resolved fluorescence microscopy [13]. Computational imaging can also be used for other purposes than for obtaining an improved resolution. For example, Light Field Microscopy (LFM) [14] provides enhanced capabilities of capturing 3D information in a single acquisition by modulation of the captured light by a microlens array. This makes refocusing of the image or choosing the direction of view possible by digital post processing only. A Whole Slide Imaging (WSI) system, the topic of this thesis, is also a digital microscope. This instrument scans a microscopic object line by line or patch by patch and merges these together in digital post processing to a large digital image. The user is presented with an image with a field-of-view far exceeding that of the optical instrument.

Computational imaging benefits from the huge progress made in computer

technology. This field has been largely driven by what is called Moore's law [15, 16], an empirical law predicting that the complexity of integrated circuits will double every two years, which has been surprisingly valid for over five decades. This has led to some developments that are of high importance for microscopy.

First, the ability of having enormous amounts of transistors on a chip has led to the development of the CMOS sensor. CMOS sensors are made in a mostly standard procedure, which allows integration of analog and digital processing on a single chip, leading to great design flexibility and miniaturization [17, 18]. The developments have been largely driven by the need for cheap, small and low power image sensors for the consumer electronics market, in particular for smartphone camera's [17, 19]. Nowadays, scientific CMOS camera's are available with high gain (on the order of single photo-electrons per digital unit), high dynamic range (up to 100 dB), very low read noise (noise equivalent of 1 photo-electron) and high frame rates (~ 100 full frames per second or ~ 400 MPx/s) [18, 20–23].

The second development is the enormous increase of storage capabilities and of distributed computing. The growth of network capacity has led to a rapid increase of what is called 'cloud' computing and storage [24]: a network based infrastructure where computer resources are delivered to users as on demand services [25, 26]. Cloud computing allows the management of unprecedented amounts of data, often called 'Big Data'.

The third development is the emergence of technologies for digital image processing. These techniques can be used for automated classification, segmentation, feature extraction or transformation of image data. The field significantly profited from the development of Graphical Processing Units (GPU's). These devices excel in executing the parallelized algorithms that are typically used in image processing. This realizes several orders of magnitude speed-up compared to calculations on a Central Processing Unit (CPU). A recent paradigm shift in digital image processing is the use of artificial intelligence techniques such as Convolutional Neural Networks (CNNs), also known as 'deep learning'. Traditional machine learning requires careful engineering to extract the desired features from the input data. In deep learning, a generic system is trained for a specific task by providing (large amounts of) example data [27]. The capability of CNNs to discover structure in high dimensional data, such as images, has led to a breakthrough in many fields of science and industry, including medicine research and the healthcare practice [28, 29], life sciences [30], and biology [31].

1.2. Whole Slide Imaging systems

A Whole Slide Imaging (WSI) system is an optical instrument aimed at creating a digital image file of a sample mounted on a microscopy slide, where the imaged area is much larger than the field of view of the microscope objective lens. The images of the WSI are typically stored on a centralized system, which delivers images on demand [32]. State-of-the-art systems allow scanning of a two-dimensional (2D) image of a 15×15 mm area at a sampling distance of $\sim 0.25\,\mu$m within a minute, resulting in image file sizes of multiple gigapixels and multiple gigabytes of compressed image data [33]. The image is presented to the user via a graphical user interface (GUI) on a computer screen that makes it possible to explore the sample at any desired location or zoom level [34].

The field of view of the optical system is extended by scanning the sample and combining multiple acquisitions digitally in post-processing, referred to as 'image stitching' or mosaicing [35]. There are two main methods of scanning [36–38]. The first is the 'step -and-stitch' method, in which the slide is moved stepwise and imaged using an area scanner. This is the most trivial extension to the traditional FOV-limited microscope but requires a potentially complex process of image-stitching in post-processing. The second is the 'push broom' or line scanning method, which combines simplicity and speed by scanning the slide with a line sensor at a constant velocity. A major challenge for efficient 2D slide scanning without the need for user interaction is tracking the topographic variations of the tissue layer, that usually exceed the depth-of-focus of the scanner. WSI systems therefore incorporate an autofocus system, either by creating a focus map of the whole slide prior to scanning [39], or by adding a closed-loop feedback system for maintaining optimum focus during scanning [40–42].

WSI systems currently on the market are developed for application in the field of digital pathology, with applications in primary diagnosis [43–46], education [47–49], and research [50, 51]. Advantages of digital pathology include the possibility of regional or even international collaboration [33], correlation with digital radiology images [52], improved workflow management and control [53], and it opens the way for Computer Aided Diagnostics (CAD) and Clinical Decision Support (CDS)[54–56].

1.3. Imaging Modalities

In this thesis, three microscopic imaging modalities are considered: brightfield imaging, fluorescence imaging and phase contrast imaging. **Brightfield imaging** is a type of microscopy in which light is passed through the sample to asses the absorption. Absorption stains are used to reveal the tissue structure, because biological material at the microscopic level is mostly transparent. The most common staining that is used is Hematoxylin and Eosin (H&E) staining [57]. Hematoxylin targets nucleic acids and gives the nuclei of the cells a deep-blue purple color. Eosin stains proteins and gives the cytoplasm and extra cellular matrix a pink color. Brightfield imaging has the advantage of being simple and excelling in revealing tissue structure [58], and is the most common modality in pathology [33].

Fluorescence microscopy images samples prepared with fluorescent stains. The physical principle of fluorescence is that a molecule is excited by light of a particular wavelength, and emits light with a longer wavelength (the so-called Stokes-shift) [59]. This difference in wavelength allows to suppress excitation light scattered by the sample towards the camera so that only the emitted fluorescence is imaged. It is probably the main microscopy modality in life sciences and biology [59, 60]. The reason is that fluorescent dyes can label specific molecules inside a cell, enabling structural biological imaging in connection to the biological function of the molecules. For example, fluorescent molecules conjugated to antibodies can be used as markers for proteins of interest, a technique known as immunohistochemistry. Also, species can be genetically modified to express fluorescent proteins, labeling the proteins of interest *in vivo*. Finally, specific genes can be labeled by using Fluorescence *in situ* Hybridisation (FISH). Another advantage of fluorescence microscopy is the ability for multiplexing. Multiple fluorophores

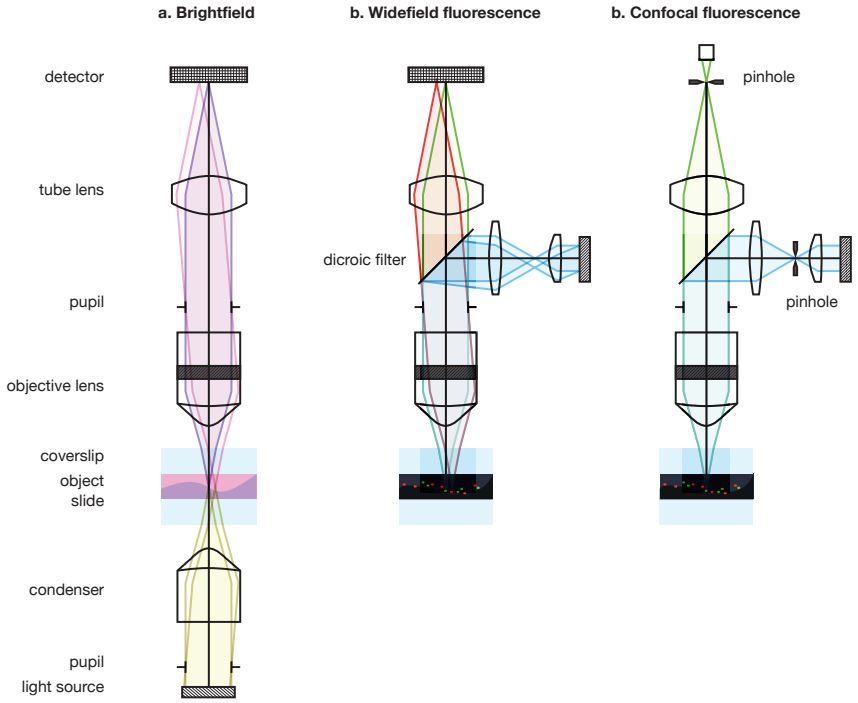


Figure 1.1: Schematic image of the optical architecture of a microscope for (a) brightfield imaging, (b) widefield fluorescence and (c) confocal fluorescence.

can be used to label different molecules of interest in the same sample. Each fluorophore can only be excited with a wavelength within a relatively narrow spectral width $\sim 10-100\,\text{nm}$ and will subsequently emit within a similarly narrow spectral width. Using light sources tuned to the excitation bandwidth (lasers or a white light source in combination with dichroic filters) can therefore selectively excite the different fluorophores. Fluorescence microscopy is implemented in a widefield imaging system (capturing the image on a camera), or in a confocal system (imaging point-by-point). In this thesis, both implementations are considered.

Quantitative phase contrast imaging reveals local variations in the optical thickness, i.e. the difference in refractive index of a sample compared to its environments. Biological structures have a natural variation in refractive index and therefore do not need staining for visualization by quantitative phase imaging. For this reason, phase contrast imaging is often used for imaging live cells [61–63]. The technique is complementary to fluorescence microscopy as it operates on unlabeled samples, has lower phototoxicity and does not suffer from photobleaching [64]. Traditional Zernike phase contrast microscopes have a modified illumination and a phase ring in the imaging path to reveal the phase contrast. With the emergence of digital microscopy, quantitative phase imaging was developed in which the phase is calculated from holographic data [64] or from a through-focus image stack obtained with a phase contrast microscope or regular brightfield mi-

croscope. Phase maps can either be 2D [65] or fully 3D [66, 67]. In this thesis we consider phase imaging using multi-focal brightfield image acquisitions.

1.4. Instrumentation

Figure 1.1 outlines the optical instruments required for brightfield, widefield fluorescence and confocal fluorescence. The optical architecture of a **brightfield imaging system** with white light illumination is shown in Fig. 1.1a. Light passes through the sample, the magnifying lens set, and is detected by an area sensor. A light source with a uniform intensity distribution across the illuminated field of view and a uniform distribution over directions of incidence is required for optimal image quality. These conditions on uniformity are achieved using Köhler illumination [68, 69]. The most straightforward method of color imaging is to use a sensor that has a Bayer filter for color sensitive pixels [70], such as is used in most digital cameras and smartphones. The disadvantages of this method are that light is lost in filtering, leading to lower signal levels, and that the sampling density for each color channel is lower than defined by the pixel pitch. Alternatively, the light can be split into a red, green and blue branch and imaged with three cameras in parallel, at the expense of an increased complexity and higher component costs.

Probably the most important system parameter is the Numerical Aperture (NA) of the objective lens, which determines the resolution that can be obtained with the system. The NA is defined as:

$$\mathrm{NA} = n \sin \alpha,$$

where n is the sample refractive index, and α the half angle of the cone of detection, within the sample. For lenses that satisfy Abbe's sine condition [69] the NA can also be expressed as:

$$\mathrm{NA} = \frac{R}{F_o},$$

where R is the radius of the pupil and F_o the focal length of the objective. The magnification of the system is given by the ratio:

$$M = \frac{F_t}{F_o},$$

where F_t is the focal length of the tube lens. The Field of View (FOV = FN/M with FN the so-called field number and M the magnification) is defined as the maximum field position that is imaged, and is ultimately limed by the objective lens. Generally, objective lenses with a high NA have a smaller FOV, due to trade-offs between the demands on optical complexity for achieving good image quality with a high NA and the corresponding demands for image quality over a large area. The ratio of the illumination NA_i and detection NA_d is called the coherence factor $\sigma = \mathrm{NA}_i/\mathrm{NA}_d$. For $\sigma \geq 1$ the system is called an incoherent imaging system and has the best resolution [69, 71].

Figure 1.1b shows a schematic view of the optical architecture of a **widefield fluorescence microscope**. This architecture is based on epi-illumination i.e., the sample is illuminated through the objective lens. The illumination light is coupled into the main optical axis by the use of a dichroic mirror. This dichroic mirror reflects light with a wavelength shorter than a cut-off wavelength (e.g. blue), while it

transmits light with a wavelength longer than the cut-off wavelength (e.g. green or red). For additional filtering, an excitation filter (not drawn) is added to the illumination arm and an emission filter (not drawn) is added to the path towards the detector [59]. A fluorescent molecule emits light isotropically, independent of the direction of excitation. This makes fluorescence microscopy incoherent, regardless of the illumination NA. There is no need for Köhler illumination in widefield fluorescence microscopy. For example, the illumination in Fig. 1.1b is placed conjugate to the pupil plane and therefore only requires to have a uniform radiation pattern for a uniform sample illumination.

The most straightforward method for multi-label imaging is to use a wideband light source, such as a mercury arc lamp, or a set of laser sources with different wavelengths, in combination with a monochromatic sensor. Then, for every fluorophore a different dichroic beam splitter is used, matched to the excitation and emission wavelength of interest [59]. The captured image data can be combined in digital post-processing, for example by assigning false colors to the different wavelengths used. Alternatively, a multi-band dichroic beam splitter might be used, which allows to divide the spectrum in multiple bands for excitation and emission. Fluorophores can then be discriminated based on their excitation wavelength, for example by switching the illumination color [59], or based on the emission spectrum, as is done in hyperspectral imaging [72–74].

Figure 1.1c shows a **confocal fluorescence microscope** architecture. The main difference with a widefield system is that here only a single point of the object is imaged at a time. A photodiode or photomultiplier is used to detect the emitted light, and a scanning approach is used to image a plane or volume inside the object. Most often, the focal point is scanned over the sample by the use of a galvo mirror system that is placed conjugate to the pupil plane (not drawn). A point source illumination is required, for example created by focusing a laser beam trough a pinhole that is optically conjugate to the object plane. Another pinhole is placed conjugate to the object plane in the detection path. As a result, light originating from an out-of-focus position is hardly illuminated, and the emitted light is to a high extent blocked by the detection pinhole. This greatly improves contrast, in particular in thick samples. This capability of optical sectioning is the primary reason of the practical success of confocal microscopy. The main system parameter in this respect is the size of the detection pinhole. A small pinhole results in an improved lateral resolution and a high background suppression, but ultimately also reduces the signal level of the in-focus fluorescence. In practice, the pinhole is chosen to have approximately the size of a focused spot. This gives a good signal level and a substantial background reduction, but does not lead to practical gains in lateral resolution. In the last decade, multiple techniques have been proposed that overcome the resolution versus signal trade-off by using a pixelated detector and re-assigning the intensity data in digital post-processing or by re-scanning the emitted light [75–79]. The point-to-point scanning strategy of confocal fluorescence microscopy results in a limited imaging throughput. Parallelization of the data acquisition is a strategy to increase speed, for example by acquiring a large number of points in parallel [80, 81].

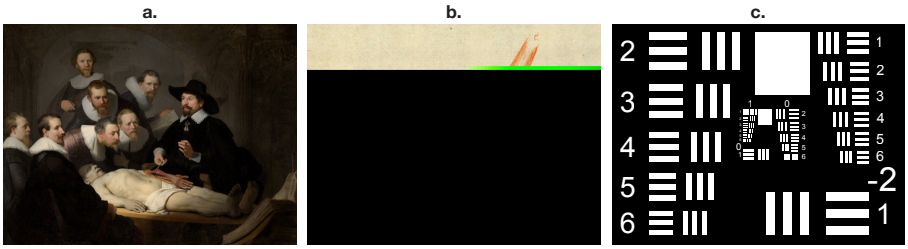


Figure 1.2: High quality images. (a) 'De anatomische les van dr. Nicolaes Tulp' (The Anatomy Lesson of dr. Nicolaes Tulp) by Rembrandt van Rijn [82]. (b) 'Een vlees-striemtge, spouwen en platdrucken' (A flesh-fiber, split and pressed flat) by Antoni van Leeuwenhoek [83]. (c) A USAF 1951 resolution test target.

1.5. Image quality

Image quality is a central theme in this thesis. Quality, however, might have many different meanings in different fields of study. Commonly, it refers to artistic quality such as the beautiful painting by Rembrandt in Fig. 1.2a showing an early study to the human body. Images of high scientific quality might refer to the object of interest, or the quality of observation such as the drawing by Van Leeuwenhoek in Fig. 1.2b, illustrating probably the first observation of muscle structure. This thesis, however, has a signal processing point of view, and regards image quality as the extent to which object features are represented in the measured signal, the image. An example is how well an optical system can resolve small structures of a resolution target, see Fig. 1.2c. In computational imaging, image quality is determined both by the imaging instrumentation and by the image processing steps that are taken to construct the final image.

Optical systems in this thesis are analysed in the framework of Fourier optics [84]. The main thesis of this framework is that a point source in the electromagnetic field in the object space of the optical imaging system will result in an extended spot in the electromagnetic field in image space, captured at the detector. This spot shape is described by the so-called Point Spread Function (PSF). The entire image can be synthesized as a sum over the PSFs corresponding to the different point sources in the imaged object. For sufficiently low NA values and for an aberration free optical system with a circular aperture the PSF is the so-called Airy pattern. This PSF has a central peaked spot with a radius $0.61\lambda/\mathrm{NA}$ and a series of low intensity concentric rings separated by dark fringes.

The PSF is closely related to the **resolution** of the optical system, probably the most important image quality metric for microscopy. Consider for example a finely spaced line grid that is imaged onto the detector. The finite sized PSF will smear out the features of the grid, ultimately making it unresolvable when the grid spacing becomes too small. The customary definition of resolution is the period of the finest line grid that can be resolved. A monumental insight in the field of microscopy was put forward by Abbe in the late 19th century [85], namely that even perfectly designed optical imaging systems have a limited resolution, and that this fundamental limit arises from diffraction at the pupil of the optical imaging system. Abbe showed that the highest spatial frequency of the object that can be de-

tected in the image is $2\text{NA}/\lambda$ (for an incoherent system), with λ the wavelength of the light, making the diffraction limit to resolution equal to $\lambda/2\text{NA}$. Note that the Numerical Aperture NA can by its definition not exceed the medium refractive index n, and as a result, the resolution is ultimately limited to the wavelength of light in the medium λ/n. Aberrations in the optical system, for example caused by imperfections in the lenses or misalignment of the components, might lead to a lower resolution than given by the diffraction limit.

Resolution and pixel pitch or 'pixel size' are often confused. From a signal processing point of view, the pixel pitch should be chosen based on the optical resolution. The Nyquist criterion states that for a digital representation of an image which is band limited with a maximum spatial frequency f_m, a sampling density $\Delta = 1/2f_m$ must be used [86]. Using a smaller pixel pitch will not improve the resolution; using a higher pixel pitch will introduce aliasing. The optical magnification is another parameter that is often confused with the resolution. Whereas in traditional microscopy the magnification played an important role for details to be visible with the human eye, this is not the case in digital microscopy where the size of the image can be chosen arbitrarily on the digital device. The optical magnification is therefore unrelated to the resolution.

Image quality is affected by **noise**. The amount of noise can be expressed by the signal-to-noise ratio (SNR), which gives the ratio between the magnitude of the signal and the standard deviation of the noise. A fundamental lower limit on the noise level follows from the particle nature of light. According to quantum mechanics light is emitted and absorbed in discrete packages called photons, where the probability of detecting a photon is proportional to the magnitude squared of the electromagnetic field at the detection point. Only photons that excite an electron in the detector actually contribute to the signal. Therefore, the measured intensity and is most often expressed in 'photo-electrons' (symbol e^-) The probability distribution for the number of photo-electrons that is measured in a certain period is given by the Poisson distribution, which has a standard deviation of $\sqrt{N}\text{e}^-$ given an expectation value of $N\text{e}^-$, an effect also referred to as shot noise. The SNR of an optical system is therefore ultimately limited to $\sqrt{N}$. Obtaining enough photo-electrons for a good SNR is challenging in low-light conditions such as in fluorescence microscopy and when short integration times are a practical constraint, such as for scanning system.

Where the fundamental limits to resolution and noise are given by the underlying physics of the optical imaging system, the image quality might in practice be limited by the **imaging sensor**. First, additional noise is usually introduced by the read-out electronics of the detector or camera. Read noise includes thermal noise from the electronic amplifiers, noise fluctuations of the dark current from the photo-diodes, reset noise from CMOS circuitry and quantization noise from the conversion of a charge or voltage to a discrete digital number [87–89]. Second, the measured value at a pixel might not only depend on the light incident at that pixel, but also by the light incident on adjacent pixels, so-called cross-talk [90], or there may be a time dependence in the read-out electronics, for example an incomplete pixel reset might introduce image lag [91]. Third, the response function of the sensor may involve unwanted complications. Typically, for microscopy applications a linear response is desired, although non-linear responses are sometimes used

to achieve a high dynamic range [90]. Linearity implies that the sensitivity of the sensor is preferably independent of the intensity it measures. Deviations from linearity can result in a false representation of the intrinsic contrast of the object. The linearity is in all cases limited by the maximum amount of photons that can be measures before the signal saturates. This full well capacity affects the dynamic range and the peak SNR of a sensor [90].

A good quality image has **uniform image features**. For this, the illumination has to be sufficiently uniform over the field of view. Also, the aberrations in the optical systems have to be well controlled to provide a uniform resolution throughout the field of view. Finally, the sensor is required to have a uniform sensitivity and offset. Pixel-to-pixel variations might lead to 'fixed-pattern-noise', degrading uniformity. In many cases, non-uniformity can be corrected for in post-processing, provided the system has a reasonable SNR.

1.6. Opportunities for WSI

The digital era brings new demands to traditional microscopy. Hardware tools are needed to generate massive amounts of image data for high throughput screening and for reliable computer aided diagnosis [92–94]. This requires imaging systems with an unlimited, or at least very large (~ cm) field of view and high resolution (~ 0.25µm). User interaction should be limited to a minimum in an environment where large amounts of samples have to be digitized at high throughput [52]. The WSI systems currently on the market for digital pathology applications satisfy these demands for 2D brightfield imaging. The goal of this thesis is to address opportunities for WSI in other imaging modalities.

First, there is a need for an efficient method for **multi-focal imaging**, also called 3D or 'z-stack' imaging [95–97]. Visualizing the 3D tissue morphology is needed in some cases in the field of histology before making a final diagnosis [36, 98]. Assessing cells in 3D is generally always needed for diagnoses in the field of cytology [52]. A multi-focal acquisition might also be used for areas in a sample where the auto-focus algorithm is not able to find focus, for example due to the presence of tissue at multiple depths.The multi-focal imaging functionality of current WSI systems is mostly realized by repeating a single layer scan for multiple positions of the objective lens. This approach is inefficient due to the considerable increase in scan time, and the sensitivity to errors in the registration of focal layers.

Second, use of WSI systems can be broaden to **quantitative phase tomography** (QPT) and 3D imaging of **thick tissue sections** (~ 60µm). Both contrast modalities are based on computational imaging techniques and are enabled by the availability of multi-focal images. Quantitative phase tomography forms a non-invasive and label-free imaging platform in cell biology and pathology [64], and can be used complementary to fluorescence imaging or for imaging of unstained tissues. 3D imaging of thick slides with widefield microscopy suffers from the lack of an optical sectioning capability, i.e. multi-focal images have optical cross-talk, adding blurred structures in out-of-focus layers to the in-focus image. This can be partly compensated by the use of a numerical post-processing technique called deconvolution [99–101]. An important requirement for both modalities is the need for simple, scalable image processing methods, ultimately enabling real time image data processing.

Third, it is a major opportunity for WSI systems to become compatible with **fluorescence imaging**. This would open digital pathology to the benefits of functional studies by immunofluorescence and FISH imaging. Both techniques offer the benefits for more specific, reproducible and quantitative diagnoses [36, 102, 103]. A WSI system capable of fluorescence imaging would also be of great use in the field of life sciences, for example for high throughput screening applications [92]. Current slide scanners are generally not capable of fluorescence imaging [104]. The few commercial solutions that are available, most often lack high throughput or do not support multiplexing.

Fourth, **confocal laser scanning microscopy** is the *de facto* standard modality for fluorescence imaging in biology [60], because the optical sectioning capability enables high contrast. The underlying point scanning technique, has a limited throughput, and the imaged area is limited by the Field Of View (FOV) of the microscope objective. This makes the technique unsuitable for fast multi-focal scanning over large areas. Potential parallelization strategies for increased throughput are the use of a line illumination instead of a spot illumination in combination with a line sensor [105–107] and scanning multiple depth layers in parallel using an illumination with multiple foci [108, 109]. Scanning the stage instead of the illumination is naturally compatible with confocal line illumination [110] and makes for a system with a minimum number of moving optical components.

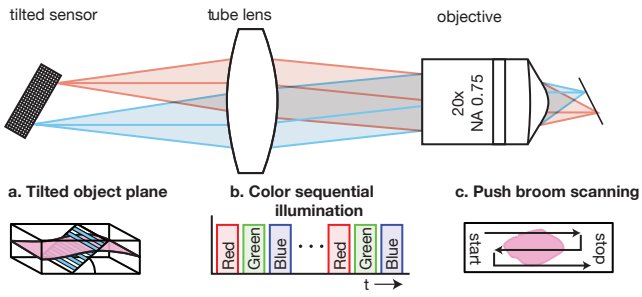


Figure 1.3: Schematic layout of the WSI system. The optical layout shows the tilted sensor that gives rise to a tilted object plane. (a) Every line on the sensor scans the sample at a different depth. (b) A color sequential illumination strategy is used for RGB imaging. (c) The slide is digitized in a push-broom scanning fashion.

The work in this thesis builds on a new WSI-system that has been devised by Philips and is described in the patent literature [111–113]. The optical architecture is illustrated in Fig. 1.3. At the heart of this architecture lies a single multi-line CMOS sensor. Placing the sensor tilted to the optical axis tilts the imaged plane in object (sample) space, see Fig. 1.3a. As a result, each sensorlet scans the tissue at a different focus level. This is used for closed loop auto-focus operation or for acquiring a multi-focal image stack in a single scan. Images with inherent color registration are acquired using a color-sequential illumination scheme [71], see Fig. 1.3b. Image acquisition is done in a 'push broom' scanning fashion [35, 114], see Fig. 1.3c, which has a limited need for stitching and results in a field of view that is in principle unlimited.

1.7. Thesis outline

This thesis contains the following chapters.

In Chapter 2 we present an in-depth analysis of the multi-line CMOS sensor for multi-focal imaging designed by Philips. This sensor allows for image data acquisition with very high throughput (~100,000 lines/s) of up to eight lines in parallel. In this chapter we describe several measurement techniques to investigate the sensor response function and the noise behavior. The main question addressed is what SNR can be obtained given a certain photo-electron flux, and the implications this has for the suitability of the sensor for fluorescence imaging.

A multi-focal imaging brightfield scanning system offers a platform for several interesting computational imaging techniques. Chapter 3 explores imaging of thick slides using an eight layer image acquisition in combination with a deconvolution technique for an improved contrast. Further, it investigates quantitative phase tomography based on a multi-focal image acquisition. An important requirement for both modalities is the need for simple, scalable image processing methods, that can easily be implemented on a parallel processing platform, ultimately enabling real time image data processing. The chapter contains a discussion of the algorithms, the numerical implementations, a detailed description of the design parameters for quantitative phase imaging, and scan results.

In Chapter 4 the extension of a WSI system to widefield fluorescence microscopy is explored. In this chapter, we limit ourselves to 2D scanning, but we use a push broom acquisition based on line-scanning compatible with a tilted sensor for 3D scanning or auto-focus. The LED based light source makes the proposed method technologically robust and cost effective. The main challenge is to provide enough optical power in the detection étendue for a substantial SNR. In this chapter we describe the design of an LED based epi-illumination module, and explore the optical power that can be delivered by an LED source, and the SNR required.

Chapter 5 investigates the use of a laser illumination source for fluorescence imaging. A laser light source is not restricted by étendue limitations, such as is the case with an LED illumination. Further, it allows to focus the light into a line that can be scanned through the sample volume and subsequently imaged onto the the detector. This results in a high illumination efficiency, and it gives an increased contrast by suppressing background light, i.e. it offers confocality. A diffractive optics method is used to generate multiple illumination lines conjugate to the tilted sensor plane. The chapter describes the design of a multi-focal (3D) scanning system, including design details of the diffractive structure, the experimental realization of a prototype system, the scan results, and a quantitative analysis of the system.

Chapter 6 concludes this thesis with a summary of the obtained results. Also, an outlook is provided to future developments in WSI systems. In particular, the challenges and some potential solutions for using a general purpose scientific CMOS (sCMOS) camera for multi-line scanning of a tilted object plane are discussed. Further, some opportunities for extension of WSI techniques to Light Sheet Microscopy (LSM) and Structured Illumination Microscopy (SIM) are considered. Finally, we discuss the opportunities for computational imaging techniques for WSI.

References

[1] R. Hooke. *Micrographia, or, some physiological descriptions of minute bodies made by magnifying glasses with observations and inquiries thereupon*. Martyn and Ja. Allestry, printers to the Royal Society, 1665. URL: https : / / archive . org / details / mobot31753000817897 (cit. on p. 3).

[2] C. K. Rosenthal. "Milestone 1 (1595) Invention of the microscope". In: *Nature Cell Biology* 11.S1 (2009), S6–S6. DOI: 10.1038/ncb1938 (cit. on p. 3).

[3] A. van Leeuwenhoek and S. Hoole. *The select works of Antony van Leeuwenhoek: containing his microscopical discoveries in many of the works of nature*. G. Sidney, 1800. URL: https://archive.org/details/selectworksofant02leeu (cit. on p. 3).

[4] A. Berk et al. *Molecular Cell Biology*. Ed. by H. Lodish. 4th ed. W. H. Freeman, 2000. ISBN: 0-7167-3136-3 (cit. on p. 3).

[5] S. Al-Janabi, A. Huisman, and P. J. Van Diest. "Digital pathology: current status and future perspectives". In: *Histopathology* 61.1 (2012), pp. 1–9. DOI: 10.1111/j.1365-2559.2011.03814.x (cit. on p. 3).

[6] *Scientist*. A google image search for 'scientist' shows a microscope in 31% of images on the first page (45 results). Interestingly, a perfect gender balance is observed with 60% of the images including one or more woman and 60% of the images including one or more man. 2019. URL: https://google.com/search?q=scientist&tbm=isch (visited on 08/29/2019) (cit. on p. 3).

[7] M. G. L. Gustafsson. "Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy". In: *Journal of Microscopy* 198.2 (2000), pp. 82–87. DOI: 10.1046/j.1365-2818.2000.00710.x (cit. on p. 4).

[8] T. Dertinger et al. *Fast, background-free, 3D super-resolution optical fluctuation imaging (SOFI)*. 2009. DOI: 10.1073/pnas.0907866106 (cit. on p. 4).

[9] S. W. Hell and J. Wichmann. "Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy". In: *Optics Letters* 19.11 (1994), p. 780. DOI: 10.1364/OL.19.000780 (cit. on p. 4).

[10] G. Patterson et al. "Superresolution Imaging using Single-Molecule Localization". In: *Annual Review of Physical Chemistry* 61.1 (2010), pp. 345–367. DOI: 10.1146/annurev.physchem.012809.103444 (cit. on p. 4).

[11] E. Betzig et al. "Imaging Intracellular Fluorescent Proteins at Nanometer Resolution". In: *Science* 313.5793 (2006), pp. 1642–1645. DOI: 10.1126/science.1127344 (cit. on p. 4).

[12] M. J. Rust, M. Bates, and X. Zhuang. "Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM)". In: *Nature Methods* 3.10 (2006), pp. 793–796. DOI: 10.1038/nmeth929 (cit. on p. 4).

[13] 'Nobel Media AB'. *The Nobel Prize in Chemistry 2014*. 2014. URL: https://www.nobelprize.org/prizes/chemistry/2014/press-release/ (visited on 11/22/2019) (cit. on p. 4).

[14] M. Levoy, Z. Zhang, and I. McDowall. "Recording and controlling the 4D light field in a microscope using microlens arrays". In: *Journal of Microscopy* 235.2 (2009), pp. 144–162. DOI: 10.1111/j.1365-2818.2009.03195.x (cit. on p. 4).

[15] G. E. Moore. "Cramming more components onto integrated circuits". In: *Electronics* 38.8 (1965), pp. 114–117 (cit. on p. 5).

[16] G. E. Moore et al. "Progress in digital integrated electronics". In: *Electron Devices Meeting*. Vol. 21. 1975, pp. 11–13 (cit. on p. 5).

[17] I. Takayanagi. "CMOS Image Sensors". In: *Image Sensors and Signal Processing for Digital Still Cameras*. Ed. by J. Nakamura. CRC Press, 2017. Chap. 5, pp. 143–178. ISBN: 9781315221083. DOI: 10.1201/9781420026856 (cit. on p. 5).

[18] A. Boukhayma. "Low-Noise CMOS Image Sensors". In: *Ultra Low Noise CMOS Image Sensors*. 2018, pp. 13–34. DOI: 10.1007/978-3-319-68774-2_2 (cit. on p. 5).

[19] R. Fontaine. “The State-of-the-Art of Mainstream CMOS Image Sensors”. In: *International Image Sensor Workshop* (2015), p. 1.01 (cit. on p. 5).

[20] Y. Wang et al. “Quantitative performance evaluation of a back-illuminated sCMOS camera with 95% QE for super-resolution localization microscopy”. In: *Cytometry Part A* 91.12 (2017), pp. 1175–1183. DOI: 10.1002/cyto.a.23282 (cit. on p. 5).

[21] S. Fullerton et al. “Optimization of precision localization microscopy using CMOS camera technology”. In: *Proc. SPIE 8228*. 2012, 82280T. DOI: 10.1117/12.906336 (cit. on p. 5).

[22] T. Maruno et al. “Comparison of CMOS and EMCCD Cameras for Computational Imaging and Application to Superresolution Localization Microscopy”. In: *Imaging and Applied Optics Technical Papers*. OSA, 2012, p. IM4C.3. ISBN: 1-55752-947-7. DOI: 10.1364/ISA.2012.IM4C.3 (cit. on p. 5).

[23] S. Fullerton et al. *Orca Flash 4.0 Changing the game*. Tech. rep. Hamamatsu Corporation, 2012 (cit. on p. 5).

[24] I. A. T. Hashem et al. “The rise of “big data” on cloud computing: Review and open research issues”. In: *Information Systems* 47 (2015), pp. 98–115. DOI: 10.1016/j.is.2014.07.006 (cit. on p. 5).

[25] R. Buyya, C. S. Yeo, and S. Venugopal. “Market-Oriented Cloud Computing: Vision, Hype, and Reality for Delivering IT Services as Computing Utilities”. In: *2008 10th IEEE International Conference on High Performance Computing and Communications*. IEEE, 2008, pp. 5–13. ISBN: 978-0-7695-3352-0. DOI: 10.1109/HPCC.2008.172 (cit. on p. 5).

[26] B. Alami Milani and N. Jafari Navimipour. “A comprehensive review of the data replication techniques in the cloud environments: Major trends and future directions”. In: *Journal of Network and Computer Applications* 64 (2016), pp. 229–238. DOI: 10.1016/j.jnca.2016.02.005 (cit. on p. 5).

[27] Y. LeCun, Y. Bengio, and G. Hinton. “Deep learning”. In: *Nature* 521.7553 (2015), pp. 436–444. DOI: 10.1038/nature14539 (cit. on p. 5).

[28] “Ascent of machine learning in medicine”. In: *Nature Materials* 18.5 (2019), pp. 407–407. DOI: 10.1038/s41563-019-0360-1 (cit. on p. 5).

[29] L. Bote-Curiel et al. “Deep Learning and Big Data in Healthcare: A Double Review for Critical Beginners”. In: *Applied Sciences* 9.11 (2019), p. 2331. DOI: 10.3390/app9112331 (cit. on p. 5).

[30] V. Pande et al. *Deep Learning for the Life Sciences*. O’Reilly Media, Inc., 2019. ISBN: 9781492039822 (cit. on p. 5).

[31] S. Webb. “Deep learning for biology”. In: *Nature* 554.7693 (2018), pp. 555–557. DOI: 10.1038/d41586-018-02174-z (cit. on p. 5).

[32] A. Huisman et al. “Creation of a fully digital pathology slide archive by high-volume tissue slide scanning”. In: *Human Pathology* 41.5 (2010), pp. 751–757. DOI: 10.1016/j.humpath.2009.08.026 (cit. on p. 5).

[33] M. D. Zarella et al. “A Practical Guide to Whole Slide Imaging: A White Paper From the Digital Pathology Association”. In: *Archives of Pathology & Laboratory Medicine* 143.2 (2019), pp. 222–234. DOI: 10.5858/arpa.2018-0343-RA (cit. on pp. 5, 6).

[34] R. A. Ruddle et al. “The Design and Evaluation of Interfaces for Navigating Gigapixel Images in Digital Pathology”. In: *ACM Transactions on Computer-Human Interaction* 23.1 (2016), pp. 1–29. DOI: 10.1145/2834117 (cit. on p. 5).

[35] S. Abeytunge et al. “Rapid confocal imaging of large areas of excised tissue with strip mosaicing”. In: *Journal of Biomedical Optics* 16.5 (2011), p. 050504. DOI: 10.1117/1.3582335 (cit. on pp. 6, 13).

[36] F. Ghaznavi et al. “Digital Imaging in Pathology: Whole-Slide Imaging and Beyond”. In: *Annual Review of Pathology: Mechanisms of Disease* 8.1 (2013), pp. 331–359. DOI: 10.1146/annurev-pathol-011811-120902 (cit. on pp. 6, 12, 13).

[37] J. R. Gilbertson et al. “Primary histologic diagnosis using automated whole slide imaging: a validation study”. In: *BMC Clinical Pathology* 6.1 (2006), p. 4. DOI: 10.1186/1472-6890-6-4 (cit. on p. 6).

[38] M. G. Rojo et al. “Critical Comparison of 31 Commercially Available Digital Slide Systems in Pathology”. In: *International Journal of Surgical Pathology* 14.4 (2006), pp. 285–305. DOI: 10.1177/1066896906292274 (cit. on p. 6).

[39] M. Montalto, R. Filkins, and R. McKay. “Autofocus methods of whole slide imaging systems and the introduction of a second-generation independent dual sensor scanning method”. In: *Journal of Pathology Informatics* 2.1 (2011), p. 44. DOI: 10.4103/2153-3539.86282 (cit. on p. 6).

[40] B. Hulsken. “Autofocus based on differential measurements”. In: *US Patent No. 9,832,365 B2* (2017) (cit. on p. 6).

[41] M. E. Bravo-Zanoguera et al. “Dynamic autofocus for continuous-scanning time-delay-and-integration image acquisition in automated microscopy”. In: *Journal of Biomedical Optics* 12.3 (2007), p. 034011. DOI: 10.1117/1.2743078 (cit. on p. 6).

[42] R.-T. Dong, U. Rashid, and J. A. Zeineh. “Dong, Rashid, Zeineh - 2005 - System and method for generating digital images of a microscope slide”. In: *US Patent Application No. 2005/0089208 A1* (2005) (cit. on p. 6).

[43] L. Pantanowitz et al. “Validating Whole Slide Imaging for Diagnostic Purposes in Pathology: Guideline from the College of American Pathologists Pathology and Laboratory Quality Center”. In: *Archives of Pathology & Laboratory Medicine* 137.12 (2013), pp. 1710–1722. DOI: 10.5858/arpa.2013-0093-CP (cit. on p. 6).

[44] R. Randell, R. A. Ruddle, and D. Treanor. “Barriers and facilitators to the introduction of digital pathology for diagnostic work.” In: *Studies in health technology and informatics* 216 (2015), pp. 443–447. DOI: 10.3233/978-1-61499-564-7-443 (cit. on p. 6).

[45] D. R. J. Snead et al. “Validation of digital pathology imaging for primary histopathological diagnosis”. In: *Histopathology* 68.7 (2016), pp. 1063–1072. DOI: 10.1111/his.12879 (cit. on p. 6).

[46] E. Abels and L. Pantanowitz. “Current state of the regulatory trajectory for whole slide imaging devices in the USA”. In: *Journal of Pathology Informatics* 8.1 (2017), p. 23. DOI: 10.4103/jpi.jpi_11_17 (cit. on p. 6).

[47] P. W. Hamilton, Y. Wang, and S. J. McCullough. “Virtual microscopy and digital pathology in training and education”. In: *APMIS* 120.4 (2012), pp. 305–315. DOI: 10.1111/j.1600-0463.2011.02869.x (cit. on p. 6).

[48] K. Foster. “Medical education in the digital age: Digital whole slide imaging as an e-learning tool”. In: *Journal of Pathology Informatics* 1.1 (2010), p. 14. DOI: 10.4103/2153-3539.68331 (cit. on p. 6).

[49] L. Pantanowitz and A. V. Parwani. “Education”. In: *Digital Pathology*. Springer International Publishing, 2016, pp. 71–78. DOI: 10.1007/978-3-319-20379-9_7 (cit. on p. 6).

[50] E. A. Krupinski, A. K. Bhattacharyya, and R. S. Weinstein. “Telepathology and Digital Pathology Research”. In: *Digital Pathology*. Springer International Publishing, 2016, pp. 41–54. DOI: 10.1007/978-3-319-20379-9_5 (cit. on p. 6).

[51] P. W. Hamilton et al. “Digital pathology and image analysis in tissue biomarker research”. In: *Methods* 70.1 (2014), pp. 59–73. DOI: 10.1016/j.ymeth.2014.06.015 (cit. on p. 6).

[52] N. Roberts et al. “Toward Routine Use of 3D Histopathology as a Research Tool”. In: *The American Journal of Pathology* 180.5 (2012), pp. 1835–1842. DOI: 10.1016/j.ajpath.2012.01.033 (cit. on pp. 6, 12).

[53] J. Griffin and D. Treanor. “Digital pathology in clinical use: where are we now and what is holding us back?” In: *Histopathology* 70.1 (2017), pp. 134–145. DOI: 10.1111/his.12993 (cit. on p. 6).

[54] M. Kowal et al. “Computer-aided diagnosis of breast cancer based on fine needle biopsy microscopic images.” In: *Computers in biology and medicine* 17.10 (2011), pp. 1563–1572 (cit. on p. 6).

[55] L. Langer et al. “Computer-aided diagnostics in digital pathology: automated evaluation of early-phase pancreatic cancer in mice”. In: *International Journal of Computer Assisted Radiology and Surgery* 10.7 (2015), pp. 1043–1054. DOI: 10.1007/s11548-014-1122-9 (cit. on p. 6).

[56] C. Castaneda et al. “Clinical decision support systems for improving diagnostic accuracy and achieving precision medicine”. In: *Journal of Clinical Bioinformatics* 5.1 (2015), p. 4. DOI: 10.1186/s13336-015-0019-3 (cit. on p. 6).

[57] J. A. Kiernan. *Histological and Histochemical Methods: Theory and Practice*. Scion, 2015. ISBN: 9781907904325 (cit. on p. 6).

[58] S. Bradbury and B. Bracegirdle. *Introduction to Light Microscopy*. Springer Singapore, 1997. ISBN: 9789813083424 (cit. on p. 6).

[59] J. W. Lichtman and J.-A. Conchello. “Fluorescence microscopy”. In: *Nature Methods* 2.12 (2005), pp. 910–919. DOI: 10.1038/nmeth817 (cit. on pp. 6, 9).

[60] J. B. Pawley. *Handbook Of Biological Confocal Microscopy*. Ed. by J. B. Pawley. Springer US, 2006. ISBN: 978-0-387-45524-2. DOI: 10.1007/978-0-387-45524-2 (cit. on pp. 6, 13).

[61] O. W. Richards. “Phase Microscopy 1954-56”. In: *Science* 124.3226 (1956), pp. 810–814. DOI: 10.1126/science.124.3226.810 (cit. on p. 7).

[62] F. Zernike. “Phase contrast, a new method for the microscopic observation of transparent objects”. In: *Physica* 9.7 (1942), pp. 686–698. DOI: 10.1016/S0031-8914(42)80035-X (cit. on p. 7).

[63] F. Zernike. “Phase contrast, a new method for the microscopic observation of transparent objects part II”. In: *Physica* 9.10 (1942), pp. 974–986. DOI: 10.1016/S0031-8914(42)80079-8 (cit. on p. 7).

[64] Y. Park, C. Depeursinge, and G. Popescu. “Quantitative phase imaging in biomedicine”. In: *Nature Photonics* 12.10 (2018), pp. 578–589. DOI: 10.1038/s41566-018-0253-x (cit. on pp. 7, 12).

[65] E. Bostan et al. “Variational Phase Imaging Using the Transport-of-Intensity Equation”. In: *IEEE Transactions on Image Processing* 25.2 (2016), pp. 807–817. DOI: 10.1109/TIP.2015.2509249 (cit. on p. 8).

[66] T. Kim et al. “White-light diffraction tomography of unlabelled live cells”. In: *Nature Photonics* 8.3 (2014), pp. 256–263. DOI: 10.1038/nphoton.2013.350 (cit. on p. 8).

[67] A. Descloux et al. “Combined multi-plane phase retrieval and super-resolution optical fluctuation imaging for 4D cell microscopy”. In: *Nature Photonics* 12.3 (2018), pp. 165–172. DOI: 10.1038/s41566-018-0109-4 (cit. on p. 8).

[68] A. Köhler. “Ein neues Beleuchtungsverfahren für mikrophotographische Zwecke”. In: *Zeitschrift für wissenschaftliche Mikroskopie und für Mikroskopische Technik* 10.4 (1893), pp. 433–440 (cit. on p. 8).

[69] M. Born and E. Wolf. *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light*. Elsevier, 2013 (cit. on p. 8).

[70] B. E. Bayer. “Color Imaging Array”. In: *US Patent No 3,971,065 A* (1975) (cit. on p. 8).

[71] S. M. Shakeri, L. J. van Vliet, and S. Stallinga. “Impact of partial coherence on the apparent optical transfer function derived from the response to amplitude edges”. In: *Applied Optics* 56.12 (2017), p. 3518. DOI: 10.1364/AO.56.003518 (cit. on pp. 8, 13).

[72] M. E. Gehm et al. “High-throughput, multiplexed pushbroom hyperspectral microscopy”. In: *Optics Express* 16.15 (2008), pp. 11032–11043. DOI: 10.1364/OE.16.011032 (cit. on p. 9).

[73] P. J. Cutler et al. “Multi-Color Quantum Dot Tracking Using a High-Speed Hyperspectral Line-Scanning Microscope”. In: *PLoS ONE* 8.5 (2013). Ed. by D. Holowka, e64320. DOI: 10.1371/journal.pone.0064320 (cit. on p. 9).

[74] H. Tsurui et al. “Seven-color Fluorescence Imaging of Tissue Samples Based on Fourier Spectroscopy and Singular Value Decomposition”. In: *Journal of Histochemistry & Cytochemistry* 48.5 (2000), pp. 653–662. DOI: 10.1177/002215540004800509 (cit. on p. 9).

[75] C. B. Müller and J. Enderlein. “Image Scanning Microscopy”. In: *Physical Review Letters* 104.19 (2010), p. 198101. DOI: 10.1103/PhysRevLett.104.198101 (cit. on p. 9).

[76] A. G. York et al. “Instant super-resolution imaging in live cells and embryos via analog image processing”. In: *Nature Methods* 10.11 (2013), pp. 1122–1126. DOI: 10.1038/nmeth.2687 (cit. on p. 9).

[77] S. Roth et al. “Optical photon reassignment microscopy (OPRA)”. In: *Optical Nanoscopy* 2.1 (2013), p. 5. DOI: 10.1186/2192-2853-2-5 (cit. on p. 9).

[78] G. M. De Luca et al. “Re-scan confocal microscopy: scanning twice for better resolution”. In: *Biomedical Optics Express* 4.11 (2013), pp. 2644–2656. DOI: 10.1364/BOE.4.002644 (cit. on p. 9).

[79] J. Huff. “The Airyscan detector from ZEISS: confocal imaging with improved signal-to-noise ratio and super-resolution”. In: *Nature Methods* 12.12 (2015), pp. i–ii. DOI: 10.1038/nmeth.f.388 (cit. on p. 9).

[80] E. Wang, C. M. Babbey, and K. W. Dunn. “Performance comparison between the high-speed Yokogawa spinning disc confocal system and single-point scanning confocal systems”. In: *Journal of Microscopy* 218.2 (2005), pp. 148–159. DOI: 10.1111/j.1365-2818.2005.01473.x (cit. on p. 9).

[81] T. Tanaami et al. “High-speed 1-frame/ms scanning confocal microscope with a microlens and Nipkow disks”. In: *Applied Optics* 41.22 (2002), pp. 4704–4708. DOI: 10.1364/AO.41.004704 (cit. on p. 9).

[82] R. van Rijn. *The Anatomy Lesson of Dr Nicolaes Tulp [Oil on canvas]*. 1632 (cit. on p. 10).

[83] A. van Leeuwenhoek. “Een vlees-striemtge, spouwen en platdrucken”. In: *Wrote Letter 35 of 1682-03-03 (AB 67) to Robert Hooke* (1682). URL: http://lensonleeuwenhoek.net/content/wrote-letter-35-1682-03-03-ab-67-robert-hooke (cit. on p. 10).

[84] J. W. Goodman. *Introduction to Fourier Optics*. 2nd ed. McGraw Hill, 1996. ISBN: 9780070242548 (cit. on p. 10).

[85] E. Abbe. “Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung”. In: *Archiv für Mikroskopische Anatomie* 9.1 (1873), pp. 413–468. DOI: 10.1007/BF02956173 (cit. on p. 10).

[86] A. V. Oppenheim, A. S. Willsky, and S. Hamid Nawab. *Signals and Systems*. 2nd ed. Pearson, 1996. ISBN: 978-0138147570 (cit. on p. 11).

[87] S. Mendis et al. “CMOS active pixel image sensors for highly integrated imaging systems”. In: *IEEE Journal of Solid-State Circuits* 32.2 (1997), pp. 187–197. DOI: 10.1109/4.551910 (cit. on p. 11).

[88] L. J. van Vliet, D. Sudar, and I. T. Young. *Digital fluorescence imaging using cooled charge-coupled device array cameras*. Ed. by J. E. Celis. 2nd ed. Vol. 3. Academic Press, 1998, pp. 109–120. ISBN: 978-0-12-164725-4 (cit. on p. 11).

[89] J. C. Mullikin et al. “Methods for CCD camera characterization”. In: *Image Acquisition and Scientific Imaging Systems*. Vol. 2173. International Society for Optics and Photonics. 1994, pp. 73–84. DOI: 10.1117/12.175165 (cit. on p. 11).

[90] W.-T. Kim et al. “A High Full Well Capacity CMOS Image Sensor for Space Applications”. In: *Sensors* 19.7 (2019), p. 1505. DOI: 10.3390/s19071505 (cit. on pp. 11, 12).

[91] Hui Tian, B. Fowler, and A. Gamal. “Analysis of temporal noise in CMOS photodiode active pixel sensor”. In: *IEEE Journal of Solid-State Circuits* 36.1 (2001), pp. 92–101. DOI: 10.1109/4.896233 (cit. on p. 11).

[92] F. Zanella, J. B. Lorens, and W. Link. “High content screening: seeing is believing”. In: *Trends in Biotechnology* 28.5 (2010), pp. 237–245. DOI: 10.1016/j.tibtech.2010.02.005 (cit. on pp. 12, 13).

[93] G. Bueno et al. “New Trends of Emerging Technologies in Digital Pathology”. In: *Pathobiology* 83.2-3 (2016), pp. 61–69. DOI: 10.1159/000443482 (cit. on p. 12).

[94] T. Dorval et al. “Filling the drug discovery gap: is high-content screening the missing link?” In: *Current Opinion in Pharmacology* 42 (2018), pp. 40–45. DOI: 10.1016/j.coph.2018.07.002 (cit. on p. 12).

[95] E. A. El-Gabry, A. V. Parwani, and L. Pantanowitz. “Whole-slide imaging: widening the scope of cytopathology”. In: *Diagnostic Histopathology* 20.12 (2014), pp. 456–461. DOI: 10.1016/j.mpdhp.2014.10.006 (cit. on p. 12).

[96] W. E. Khalbuss, L. Pantanowitz, and A. V. Parwani. “Digital Imaging in Cytopathology”. In: *Pathology Research International* 2011 (2011), pp. 1–10. DOI: 10.4061/2011/264683 (cit. on p. 12).

[97] T. Kalinski et al. “Virtual 3D Microscopy Using Multiplane Whole Slide Images in Diagnostic Pathology”. In: *American Journal of Clinical Pathology* 130.2 (2008), pp. 259–264. DOI: 10.1309/QAM22Y85QCV5JM47 (cit. on p. 12).

[98] T. C. Cornish, R. E. Swapp, and K. J. Kaplan. “Whole-slide Imaging: Routine Pathologic Diagnosis”. In: *Advances In Anatomic Pathology* 19.3 (2012), pp. 152–159. DOI: 10.1097/PAP.0b013e318253459e (cit. on p. 12).

[99] G. M. P. van Kempen et al. “A quantitative comparison of image restoration methods for confocal microscopy”. In: *Journal of Microscopy* 185.3 (1997), pp. 354–365. DOI: 10.1046/j.1365-2818.1997.d01-629.x (cit. on p. 12).

[100] R. Heintzmann. “Estimating missing information by maximum likelihood deconvolution”. In: *Micron* 38.2 (2007), pp. 136–144. DOI: 10.1016/j.micron.2006.07.009 (cit. on p. 12).

[101] J. B. Sibarita. “Deconvolution microscopy”. In: *Advances in Biochemical Engineering/Biotechnology* 95 (2005), pp. 201–243. DOI: 10.1007/b102215 (cit. on p. 12).

[102] C. Higgins. “Applications and challenges of digital pathology and whole slide imaging”. In: *Biotechnic & Histochemistry* 90.5 (2015), pp. 341–347. DOI: 10.3109/10520295.2015.1044566 (cit. on p. 13).

[103] H.-A. Lehr et al. “Quantitative Evaluation of HER-2/ neu Status in Breast Cancer by Fluorescence In Situ Hybridization and by Immunohistochemistry With Image Analysis”. In: *American Journal of Clinical Pathology* 115.6 (2001), pp. 814–822. DOI: 10.1309/AJ84-50AK-1X1B-1Q4C (cit. on p. 13).

[104] L. Pantanowitz, N. Farahani, and A. Parwani. “Whole slide imaging in pathology: advantages, limitations, and emerging perspectives”. In: *Pathology and Laboratory Medicine International* 7 (2015), p. 23. DOI: 10.2147/PLMI.S59826 (cit. on p. 13).

[105] L. Li et al. “High-throughput imaging: Focusing in on drug discovery in 3D”. In: *Methods* 96 (2016), pp. 97–102. DOI: 10.1016/j.ymeth.2015.11.013 (cit. on p. 13).

[106] R. Wolleschensky, B. Zimmermann, and M. Kempe. “High-speed confocal fluorescence imaging with a novel line scanning microscope”. In: *Journal of Biomedical Optics* 11.6 (2006), p. 064011. DOI: 10.1117/1.2402110 (cit. on p. 13).

[107] K.-b. Im et al. “Simple high-speed confocal line-scanning microscope”. In: *Optics Express* 13.13 (2005), pp. 5151–5156. DOI: 10.1364/OPEX.13.005151 (cit. on p. 13).

[108] P.-H. Wang et al. “Non-axial-scanning multifocal confocal microscopy with multiplexed volume holographic gratings”. In: *Optics Letters* 42.2 (2017), pp. 346–349. DOI: 10.1364/OL.42.000346 (cit. on p. 13).

[109] A. Jesacher, S. Bernet, and M. Ritsch-Marte. “Colored point spread function engineering for parallel confocal microscopy”. In: *Optics Express* 24.24 (2016), pp. 27395–27402. DOI: 10.1364/OE.24.027395 (cit. on p. 13).

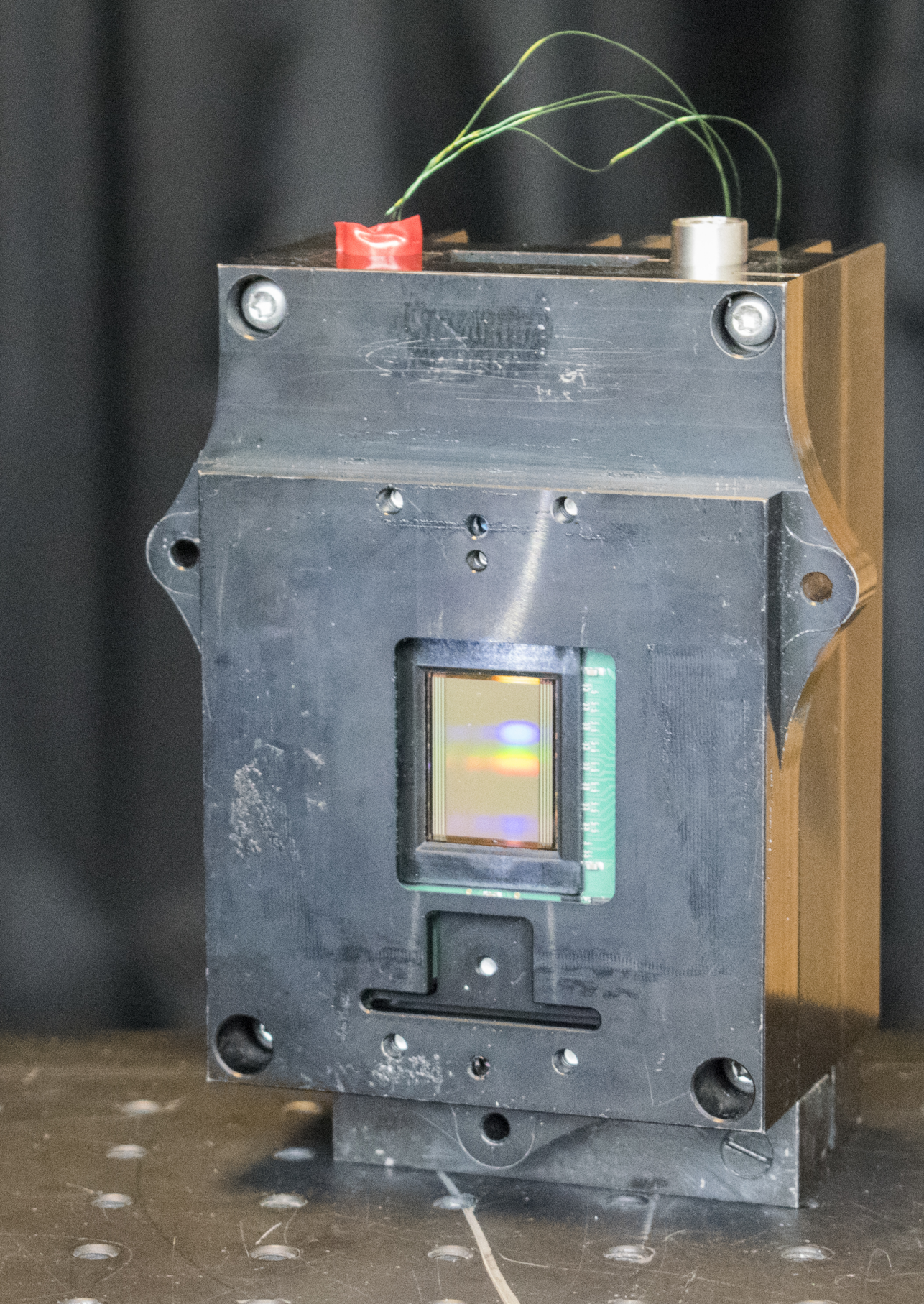
[110] T. Yang et al. “Rapid imaging of large tissues using high-resolution stage-scanning microscopy”. In: *Biomedical Optics Express* 6.5 (2015), pp. 1867–1875. DOI: 10.1364/BOE.6.001867 (cit. on p. 13).

[111] B. Hulsken. “Method for simultaneous capture of image data at multiple depths of a sample”. In: *US Patent No. 9,910,258 B2* (2018) (cit. on p. 13).

[112] B. Hulsken. “Scanning imaging system with a novel imaging sensor with gaps for electronic circuitry”. In: *US Patent. No. 10,091,445 B2* (2018) (cit. on p. 13).

[113] B. Hulsken and S. Stallinga. “Sensor for microscopy”. In: *US Patent No. 10,353,190 B2* (2019) (cit. on p. 13).

[114] S. M. Shakeri et al. “Optical quality assessment of whole slide imaging systems for digital pathology”. In: *Optics Express* 23.2 (2015), pp. 1319–1336. DOI: 10.1364/OE.23.001319 (cit. on p. 13).



2

Experimental characterization of a novel multi-line image sensor for application in whole slide imaging

A novel architecture for whole slide imaging with simultaneous image acquisition and auto-focus imaging at very high speed (>100 Mpixel/s) is investigated. The architecture is based on a custom designed CMOS sensor that consists of multiple sensorlets that can be individually selected for readout. The sensor is placed at an angle with the optical axis, so that each sensorlet can generate a scan plane at a different depth in the sample. We performed an experimental characterization to asses the sensor for image artifacts and noise behavior. The sensor is found to show a non-linearity in the readout leading to a discontinuity in the histogram of digital pixel values. A mitigation to this 'missing symbols' artifact is proposed in the form of a non-linear mapping that can be applied in digital post processing. Also, the sensor is found to show high frequency, pixel value dependent, variations in the gain, leading to a striping artifact in the brighter areas of the image, oriented perpendicular to the scan direction. Additionally, minor artifacts were identified including a fluctuating background and image lag. For these artifacts, straightforward linear unmixing type corrections are available. A rather high read-out noise of about 40 photoelectrons is found for a wide range of gain settings. This has only a small impact on brightfield imaging, where a large photon current is typically achieved, and for which a very high signal-to-noise ratio (SNR) of 292 can be obtained. It does pose a problem for fluorescence imaging, however, where the signal photon flux is much lower, which compromises the SNR. We show that fluorescence imaging is not possible using LED illumination and only at a limited speed of around 1-10 Mpixel/s using laser illumination.

2.1. Introduction

The optical microscope is one of the core instruments used by pathologists to study tissues at length scales down to the cellular level for diagnostic purposes. The advances in image digitizing technologies have led to the field of 'digital pathology', in which the traditional microscope is replaced by a Whole Slide Imaging (WSI) system. These systems enable scanning of tissue slides at high resolution (~0.25 µm) and high throughput (15 mm^2/min) to digital image files [1, 2]. Different scanning approaches are available, of which the line-sensor based 'push broom' approach for two-dimensional (2D) brightfield imaging seems prevalent [3]. There are several challenges in whole slide imaging. First, accurate focusing for 2D imaging without any user interaction practically involved is a critical challenge, primarily due to inherent tissue topography variability [4–6]. Second, 2D imaging does not cover all use cases. It is regularly required to asses the cellular composition of specimens in three dimensions in the field of histopathology and cytology [7–10]. Three-dimensional (3D) scanning in current WSI systems is mostly realized by scanning a series of adjacent tissue sections or through multiple scans at different focus levels to create a so-called '3D virtual slide' or 'z-stack'. This approach is sensitive to misalignment and might need a post-processing registration step. Also, the scan duration is increased considerably and consequently hampers adoption in clinical practice.

A conceptually novel 3D WSI architecture based on a custom-built multi-line CMOS image sensor has been proposed by Philips[11–13]. The sensor contains 128 sensorlets of 4×4096 pixels. Placing the sensor tilted to the optical axis tilts the imaged plane in object (sample) space. As a result, each sensorlet scans the tissue at a different focus level. This can be used for closed loop auto-focus operation or for acquiring a 3D image stack in a single scan.

One of the main advantages of CMOS technology is the integration of the image sensor with circuitry for both driving the image sensor and performing on-chip signal processing [14]. In addition, CMOS technology allows for very high read-out speeds in comparison to CCD, where the charge transfer read-out requires a serial read-out. A disadvantage of CMOS technology is the reduced photon detection efficiency due to the requirement of multiple transistors per pixel, which occupy space on the sensor surface [14, 15]. Also, CMOS sensors suffer from fixed pattern noise (FPN), which is caused by pixel-to-pixel variations due to device and interconnect mismatches across the sensor array, and from reset noise that is caused by an incomplete pixel reset [15, 16]. Offset FPN and reset noise can be reduced by correlated double sampling (CDS) at the expense of more transistors per pixel [15]. The low photon detection efficiency can be increased by further miniaturization of the electronics and the use of micro-lenses [15, 17]. The novel designed multi-line CMOS image sensor makes use of the advantages of CMOS technology by having integrated read-out electronics and two Field Programmable Gate Arrays (FPGAs) for on-chip processing of the image data. The read-out electronics is relocated to the regions of the sensor between the sensorlets that are not needed as photosensitive regions. This increases the fill factor to 79 % and because in total more circuitry can be positioned on the sensor, enables a faster read-out. It also eliminates the need for micro-lenses, which are undesirable when the image sensor is tilted with respect to the optical path [12]. The issue of FPN is addressed by the inclusion of a

series of 'dark pixels' on every row on the image sensor. This allows for removal of FPN in numerical post processing.

Here, we report on an experimental characterization of the custom-built multi-line CMOS image sensor. The first goal is to test the sensor for image artifacts and to assess the possibility of correcting these artifacts in digital post processing. The second goal is to determine the signal-to-noise ratio (SNR) of the system as a function of the available photo-electron current, based on an experimental evaluation of different noise sources. In our experimental evaluation and characterization we build on methods that have been reported in the literature. In particular, we make use of the 'photon transfer method', which exploits the linear relation of the shot noise induced variance with the number of photoelectrons in order to measure gain and read noise [18–20]. We also make use of existing experimental methods to measure dark current and linearity [19, 21, 22]. Recently, we investigated the possibilities and limits of extending WSI systems to fluorescence imaging by using a LED based epi-illumination [23]. Typical photo-currents that can be achieved for fluorescence imaging are orders of magnitude smaller than for brightfield imaging. For that reason we will focus specifically on the sensor performance at low light levels.

This report will start with a brief description of the novel 3D optical architecture and with the sensors capabilities in Section 2.2. Then, we will give an effective sensor model describing the stochastic process of data acquisition in Section 2.3. We have characterized a number of effects and artifacts, and measured performance parameters. The different measurement methods as well as the key results are reported in Section 2.4. We conclude this chapter with a discussion of the implication of our measurements for the achievable SNR, the consequences for fluorescence scanning and the next steps of research into sensor performance in Section 2.5.

2.2. Whole slide scanner architecture

Slide scanning is most often implemented by using a line scanning setup with continuous scanning ('push broom' scanning). This is a favored architecture for its mechanical simplicity and reduced need for stitching [24]. Figure 2.1a shows a novel line scanner architecture based on an image sensor that is tilted with respect to the optical axis. This gives rise to a tilted object plane, as illustrated in Fig. 2.1b. Therefore, every row on the sensor corresponds to a different z-position in object space. Scanning at multiple depth levels can be used for auto-focus or for multi-focus volumetric (3D) imaging. A time-sequential illumination strategy is used (see Fig. 2.1b) to acquire color images (up to five channels, but default three for RGB). This scanning architecture removes the need for multiple sensors for auto-focus and color and gives intrinsic alignment of the color channels and depth layers.

The used multi-line sensor is a custom designed CMOS image sensor which can capture high resolution images and auto-focus images simultaneously at a very high lines rate up to 100 klines/s. The design of the image sensor originates from Philips and has been described in the patent literature [11–13]. The pixel geometry of the sensor is shown in Fig. 2.1d. The sensor consists of 128 sensorlets, each sensorlet consisting of four rows. The sensor is 4096 pixels wide. In addition,

32 extra pixels are placed on both sides which are not sensitive to light, with the purpose of background correction. The pixel pitch is 6.42 µm in the scan direction and 5.54 µm in the lateral direction. The projection of the tilted rectangular pixels onto the scanned plane is square for a sensor tilt of about 30°.

The sensor unit comprises two Field Programmable Gate Array (FPGA) modules or 'engines' for reading and processing the image data. Both engines have four 9-bit Analog-to-Digital-Convertors (ADCs). Each ADC can be independently configured to read a line of a sensorlet. The first ADC will always read the first line, the second ADC the second line, etc.

The main purpose of having two engines is to use one engine for acquiring high resolution image data, while the other engine is used to capture auto-focus data. This is illustrated in Fig. 2.1e. Capturing high resolution images can make use of digital Time Delayed Integration [25] (TDI): the four ADCs of an engine read the four rows of a single sensorlet, and the data is added with a delay in order to achieve a fourfold increase in signal. Simultaneously, the other engine reads out a series of sensorlets sequentially. This is most often done in a 'binning' mode, where two ADCs read two lines of a sensorlet and donwsample the data with a factor two for SNR improvement. The third and fourth ADC read the next sensorlet simultaneously. For the next line, the next two sensorlets are read. This is repeated until for every sensorlet a line of downsampled data is available. This dataset is used to determine optimal focus and adjust the position of the microscope objective. However, capturing auto-focus data is not the only possibility this sensor architecture offers. Alternatively, the engines can be used to simultaneously acquire data from eight different sensorlets, as illustrated in Fig. 2.1f. The result is a multi-focal volume or 'z-stack' of image data.

The software interface (provided by Philips) supports multiple predefined acquisition modes and settings. Table 2.1 lists the most important options. There are four Sensor modes listed. 1xTDIx1 is the most simple mode, in which one ADC is used to read one row of the selected sensorlet, resulting in one image. In the 1xTDIx4 mode, four images are acquired: all ADCs of an engine are set to read a different sensorlet. In the 4xTDIx1 mode, all ADCs are used to read the four lines of a single sensorlet. The engine performs online TDI and outputs a single image. 2xTDIx2 is a mixed mode in which two sensorlets are selected for double-row TDI. The binning mode is designed for auto focusing and most often combined with a scanning row address mode. It works comparable to the 2xTDIx2 mode but performs a factor two downsampling. The row address mode is either in fixed or 'scanning' mode. In scanning mode, the selected sensorlet is changed every line. The start sensorlet, step size and number of steps can be defined.

2.3. Effective sensor model

A digital CMOS sensor is an electronic device that converts collected photoelectrons to a voltage, which is amplified and subsequently converted to a digital number. In this section we present an effective sensor model to describe the stochastic process of acquiring the value X of a single pixel. We start with the general description

$$X = \mathscr{D}\left[\frac{Q+B+N}{g}\right], \tag{2.1}$$

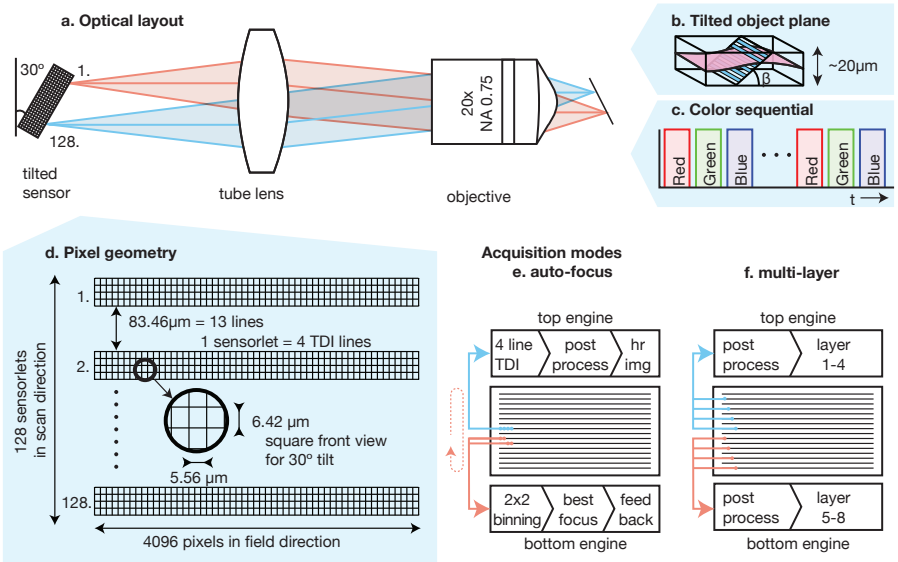


Figure 2.1: Illustration of the scanning system.
a. Schematic layout of the WSI system, showing the rotated sensor, tube lens and the tilted object plane. Depending on the chosen sensorlet (e.g. 1 in red or 128 in blue) a different axial position is scanned.
b. The sensor is tilted with respect to the optical axis. This make the sensorlets correspond to different z-positions in the objects space. For a sensor tilted 30° and using an 20× objective, a range of 20 µm is covered.
c. The sensor is monochromatic. Color images are obtained by using a synchronized light source that sequentially illuminates the sample with red green and blue to obtain an RGB image.
d. The sensor consists of 128 sensorlets of four rows which are 4096 pixels wide. The space between the sensorlets equals 13 rows. The pixel size is 6.42 µm in the scan direction and 5.54 µm in the lateral direction.
The sensor contains two 'engines' (top and bottom) which both can read-out four lines simultaneously. The diagram shows the connections of the sensorlets to the top and bottom engine in different acquisition modes.
e) For 2D scanning, the top engine is used to read all four rows of a selected sensorlet and preform four row TDI to obtain a high resolution image. The bottom engine is used in binning mode to read and downsample the data of two rows of a sensorlet. Two sensorlets can be read-out simultaneously. Every timestep, the next two sensorlets are read, such that after 64 steps, all sensorlets have been read and the optimal focus can be calculated to provide closed loop feedback for auto-focus.
f) For 3D scanning, both engines read four sensorlets to obtain an 8 layer multi-focal volume

Table 2.1: List of relevant options in the Philips sensor software.

Setting	name	options	can be set per
Sensor mode	SensorMode	1xTDIx1, 4xTDIx1, 1xTDIx4, 2xTDIx2 Binning	engine
Row address mode	RowAddrMode	Fixed, Scanning	engine
Analog Gain	AnalogGain	0-255[1] (default = 32)	engine
Analog Offset	AnalogOffset	0-255[1]	engine
Number of lines	NrLineTriggers	integer	scan
Line period	LineTriggerPeriod	35.714 µs (default) to 892.85 µs (25× default)[2]	scan
Exposure time	ExposureTimeCh[0,1,2]	8.432 µs (default) to 210.8 µs (25× default)[2]	channel

1. Throughout this document a decimal notation will be used for these settings, opposed to the software interface that uses a hexadecimal notation.
2. Actual limits might be slightly different. The mentioned range was used for this research.

where $\mathscr{D}$ denotes the conversion to a digital number, Q the collected charge in e^-, B the electron equivalent background offset, N the electron equivalent read noise, and g the gain in e^-/ADU.

The charge Q is the accumulation of the current that is created by electrons that are excited in the semiconductor. Light falling on the sensor gives rise to a current I that is proportional to the intensity. Additionally, thermal excitation of electrons in the semiconductor gives rise to a 'dark current' I_d. Both currents are accumulated during the exposure time τ. Due to the discrete nature of the electrons, both processes are stochastic and can be described using Poisson statistics i.e. they give rise to shot noise. The sensor has a built-in dark current compensation. Therefore, the dark current will effectively not increase the collected charge, but will contribute to the variance of the collected charge. A common behavior of a CMOS sensor is to show some image lag, which is caused by an incomplete pixel reset [16]. That means that a fraction l of the charge of the previous acquisition Q_{-1} is remaining. In summary,

$$\mathrm{E}[Q] = I\tau + lQ_{-1}, \tag{2.2}$$

$$\mathrm{var}[Q] = (I + I_d)\,\tau + l^2\mathrm{var}[Q_{-1}]. \tag{2.3}$$

The background profile B in a CMOS sensor is the result of the offset voltage of the analog amplifier. In the sensor under investigation, this offset can be adjusted using the Analog Offset setting. The offset is depending on the pixel position y and the chosen ADC. Additionally, this offset shows a position independent variation in time. As this can be fully corrected for using the black pixels, this is not treated as noise. The background profile is written as the sum of a static and a fluctuating

part.

$$B = B_s(y) + B_f(t) \tag{2.4}$$

The read noise N is a collection term for all temporal noise in the process of reading a pixel [14, 19]. The underlying physical processes include noise in the read-out electronics, and variations in the pixel reset [14, 16]. The read noise depends on the configuration of the analog amplifier and will therefore depend on the Analog Gain setting. Read noise has zero mean and a variance

$$\mathrm{var}[N] = \sigma_0^2. \tag{2.5}$$

Before conversion to a digital number, the voltage is amplified to match the dynamic range of the ADC, which determines the gain g. This amplification can be configured using the Analog Gain setting of the sensor. In our sensor, the gain varies between the ADCs. Also, the amplification shows a non-linear, time dependent fluctuation. Therefore, we model

$$g = g_0 + g_f(Q, t), \tag{2.6}$$

where g_0 is the actual sensor gain for the given ADC expressed in e^-/ADU. The second term involving g_f represents the amplification fluctuations, where by definition $\mathrm{E}[g_f] = 0$.

The analog to digital conversion performs the mapping of a range of voltages to digital numbers. In this conversion a round-off error or 'quantization noise' is introduced which can be approximated as an additive white noise source with a variance of $(1/12)\mathrm{ADU}^2$ [19]. Ideally, the ADC performs uniform mapping of voltage to digital values. However, under certain conditions, we found the sensor to do a non-linear mapping of voltages to digital numbers. Therefore, we write

$$\mathscr{D}[X] = M(\mathscr{D}_0[X]), \tag{2.7}$$

where $\mathscr{D}_0[X]$ represents regular analog to digital conversion. $M(X)$ represents the effect of a non linear mapping relating the input and output in ADU. If this mapping is a unique mapping, there exists a function M^{-1} such that $X = M^{-1}(M(X))$, which implies that this effect can be numerically corrected. The non-linear mapping in the sensor under investigation is a result of an effect that will be called 'missing symbols'. Summarizing all these definitions, we find

$$X = M\left(\mathscr{D}_0\left[\frac{(I + I_d)\tau + lQ_{-1} + B_s(y) + B_f(t) + N}{g + g_0(t, Q)}\right]\right) \tag{2.8}$$

We distinguish three type of effects: those that are invertible: image lag and background, those that cause non-linearities: missing symbols and gain fluctuations and those that contribute to noise: dark current, quantization, and read noise. The signal-to-noise ratio (SNR) is given by:

$$\mathrm{SNR} = \frac{I\tau}{(I + I_d)\tau + \sigma_0^2 + \frac{1}{12}g_0^2} \equiv \frac{I\tau}{I\tau + \sigma}, \tag{2.9}$$

where, for convenience, we defined σ to represent the total amount of signal independent noise of the system.

2.4. Sensor characterization

2.4.1. Experimental setup

A whole slide scanner prototype is built for testing the novel multi line sensor. We use a Nikon 20× NA 0.75 Plan Apochromat VC objective lens and a custom Nikon tube lens (effective focal length of 222.4 ± 2.2 mm) for obtaining a target magnification of $M = 22.24$. A sample stage is made using two stages: a PI M-505 low profile translation stage is used for positioning of the slide in the field direction and a Newport XM1000 ultra precision linear motor stage for the continues scanning motion of the sample. To focus, we use two stages to axially translate the objective: a PI M-111 compact micro-translation stage for coarse positioning, and a PI P-721.CL0 piezo nano-positioner for fine positioning. We use a custom RGB LED based light source for widefield illumination. The custom made LED driver has closed loop optical feedback for power stabilization and switching times on the order of 2 µs [23]. For our measurements, we use the blue channel, which contains a Lumiled Luxeon Rebel Blue (~470 nm, LXML-PB02) LED. The illumination is controlled by a National Instruments NI PXIe-6363 data acquisition card and is synchronized to the sensor.

2.4.2. Notation conventions

The data of an engine of the sensor is described as as a four dimensional digital dataset. This is illustrated in Fig. 2.2. The notation will be used that x_{lpcz} denotes a measured value of line l, pixel p, channel c and layer z. The dataset has $N_l \times N_p \times N_c \times N_z$ elements. The fist dimension is the 'scan' direction and has a size N_l equal to the Number of lines setting. The second dimension is the lateral direction with $N_p = 4096$ pixels. The third dimension represents the color channels, with $N_c \leq 5$ and which has $N_c = 3$ for RGB imaging. Lastly, the dataset has N_z layers. Lines and channels are acquired line after line, and for every line N_c channels are acquired sequentially. The connotation of the term 'line' is therefore temporal in the current context and must not be confused with a physical row of pixels on the sensor. Every layer corresponds to a physical sensorlet. In our measurements, the sensor is only used in a 'Fixed' Row address mode. This means that the data of all layers is captured simultaneously and N_z is solely dependent on the Sensor mode setting, e.g. $N_z = 1$ for 1×TDI×1 and 1×TDI×4, but $N_z = 4$ for 4×TDI×1. The dark current pixels are denoted by $b_{lp'cz}$ with corresponding dimensions, except for the second dimension which has $N_{p'} = 64$ corresponding to the left 32 and right 32 black pixels respectively. In most cases we will work with the data of a single layer and a single channel. The last two indices will then be left out for brevity.

We have characterized a number of effects and artifacts, and measured performance parameters. These are, missing symbols, background pattern, image lag, dark current, fluctuations of gain, the sensor gain as a function of the Analog Gain setting and the ADC, and the linearity of response. The measurement methods as well as the key results are summarized in the following.

2.4.3. Missing symbols

The sensor is tested for a non-linear mapping of voltages to digital numbers, with a possible 'missing symbols' like discontinuity, by testing the sensor response in

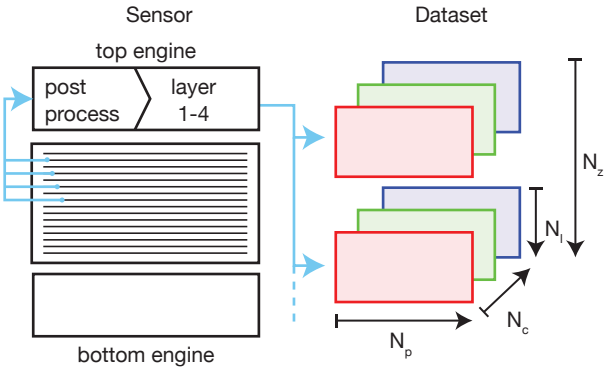


Figure 2.2: Illustration of the notation convention of the four dimensional dataset generated by the sensor.

the scan direction. The sensor is illuminated uniformly. The light source is pulsed in synchronization with the sensor integration period. The pulse sequence is configured to illuminate only one channel. The pulse duration is increased from zero to the full exposure time during the course of a scan to create a gradually changing illumination spanning the full dynamic range of the sensor. It is anticipated that, although the illumination power increases smoothly, the effect of missing symbols will cause discontinuities in the obtained data. The result is demonstrated in Figs. 2.4a to 2.4c which displays the value of pixel 2000 as a function of time for all three color channels, respectively. The measured value is increasing with the line as expected. However, a range of values is clearly missing, indicating a non-linear response with a discontinuity. For the first (red) and second (green) channel, this occurs at about 120 ADU and for the third (blue) channel around 300 ADU. Not only pixel 2000 is affected, but all pixels are missing the same range of values. This can be seen in the histograms in Figs. 2.4d to 2.4f that are made including all pixels.

To gain further understanding of the actual behavior of the missing symbols, we measured a static defocused resolution target. The raw data is shown in Fig. 2.5a and shows a sharp edge between the background and the brighter areas. Plotting the image data in a histogram, given in Fig. 2.5c, clearly shows a range of missing symbols. Fig. 2.5e shows the pixel variance v_p as a function of the pixel mean m_p, where the variance and mean of each pixel are computed over the set of lines acquired sequentially over time, i.e. they are defined as:

$$m_p = \text{mean}_l[x_{lp}] \tag{2.10}$$

$$v_p = \text{var}_l[x_{lp}] \tag{2.11}$$

An increased variance is present around the range of missing values. Just above this range, a decrease of the variance is observed.

We propose the following piecewise linear model for the mapping function $M(X)$ to explain the observed behavior.

$$M(X) = \begin{cases} X + d\frac{X-X_2}{X_2-X_1}, & \text{where } X_1 < X < X_2 \\ X & \text{otherwise} \end{cases} \tag{2.12}$$

In this model, a discontinuity occurs at X_1 and has size d ADU which is added to the output. Over the range X_1 to X_2 this value decreases linearly to zero. The model

is illustrated in Fig. 2.3a. The discontinuity in the model explains the increase in variance, which is caused by small deviations that are amplified by the discontinuity in $M(X)$ giving rise to an added variance on the order of d^2. This range of lower variance is explained by the region between X_1 and X_2 of the mapping function $M(X)$ which has effectively a lower gain. In case that $d < X_2 - X_1$ the mapping between input and output remains unique and therefore, invertible:

$$M^{-1}(Y) = \begin{cases} X_1 & X_1 \leq Y < X_1 + d \\ Y - d\frac{Y - X_2}{X_1 + d - X_2} & X_1 + d \leq Y < X_2 \\ Y & \text{otherwise} \end{cases}. \tag{2.13}$$

Here, whe chose to map values between X_1 and $X_1 + d$, which are not expected to occur, to X_1. The inverse mapping is illustrated in Fig. 2.3b.

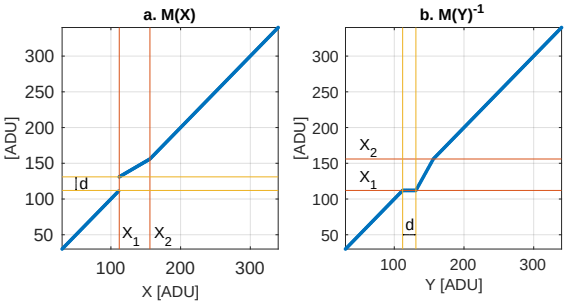


Figure 2.3: Illustration of the proposed model for the non-linear mapping function.
(a) The forward mapping $M(X)$, with indicated the start of the non-linear region X_1, the size of the discontinuity d and the end of the non-linear region X_2.
(b) The inverse mapping function $M^{-1}(Y)$.

The model parameters X_1, X_2, and d introduced in Eq. (2.12) are found by a numerical inversion scheme. First, given some values of X_1, X_2 and d, the corrected value x^c is calculated using

$$x^c = M^{-1}(x), \tag{2.14}$$

where the superscript c indicates the correction for missing symbols. From the corrected values, the corrected mean $m_p^c = \text{mean}_l[x_{lp}^c]$ and corrected variance $v_p^c = \text{var}_l[x_{lp}^c]$ are calculated. By repeating this procedure for all possible values of X_1, X_2 and d the overall deviation of v_p^c from a linear relation with m_p^c is minimized:

$$\underset{X_1 X_2 d}{\arg\min} \sum_{p=1}^{N_p} \left[v_p^c - (A + B m_p^c) \right]^2, \tag{2.15}$$

where A and B are the least squares solution of the linear model $v_p^c = A + B m_p^c$. For the given sensor settings, the values $X_1 = 112\,\text{ADU}$, $X_2 = 156\,\text{ADU}$, and $d = 19\,\text{ADU}$ are found. In the right column of Fig. 2.5, the correction of missing symbols using the inverse function M^{-1} given in Eq. (2.13) is demonstrated. Figure 2.5b shows the image data after correction. The inverse function indeed removes the gap in the histogram, as is show in Fig. 2.5d and the deviations in the variance as shown in Fig. 2.5f.

The range of values at which the missing symbols occur depends on the exposure time. A measurement is performed for a range of exposure times between

8.432 µs and 25.296 µs (1 and 3× default), see Fig. 2.6. The range of missing values shifts proportionally to higher values for increasing exposure times. Extrapolating the results, it reaches the end of the dynamic range at about 34 µs (4× default).

The amount of values missing also depends on the Analog Gain, as can be seen in the histograms in Fig. 2.7. The range of affected values changes from 15 ADU at an Analog Gain of 6 to 3 ADU at an Analog Gain of 50. Using the sensor gain, see Section 2.4.8, it is calculated that this corresponds to roughly 200 e^- regardless of the chosen setting.

The effect of missing symbols is not caused by interference between sequential lines, or from an incomplete pixel reset. This is verified by repeating the previous experiment, but now with an illumination sequence in which only one out of three lines is exposed. This measurement shows the same behavior. Additionally, it was found that the Analog Offset setting does not change the range of missing values.

Our results imply that lower line rates can be used as a work around. Throughout the measurements in the remainder of this section, the sensor is operated such that the effect of missing symbols is avoided.

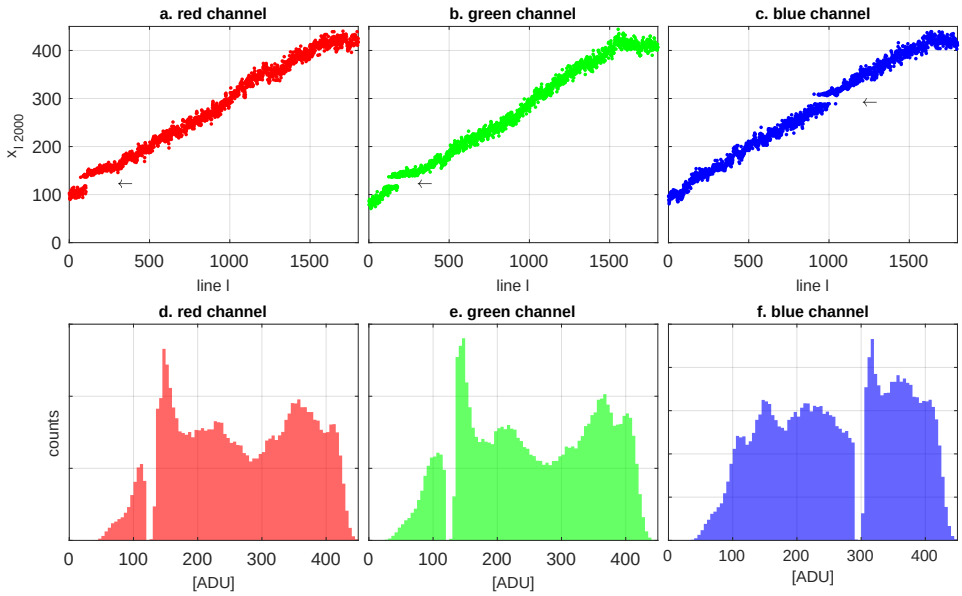


Figure 2.4: Demonstration of missing symbols. The sensor was operated in 1xTDIx1 mode using sensorlet 64, an Analog Gain of 6 and an Analog Offset of 113.
a-c) Graphs showing the value of pixel 2000 as a function of time for the red, green and blue channel.
d-f) Histogram of the full scan, showing missing symbols at the same absolute ADU values for the red, green and blue channel.

2.4.4. Image lag

The sensor is tested for image lag in a bright-to-dark transition. The sensor is uniformly illuminated with a pulse sequence in which only one channel of one out of three lines is exposed. Over the full course of the measurement, the pulse length of the illumination is increased from zero to about two times the exposure time. The

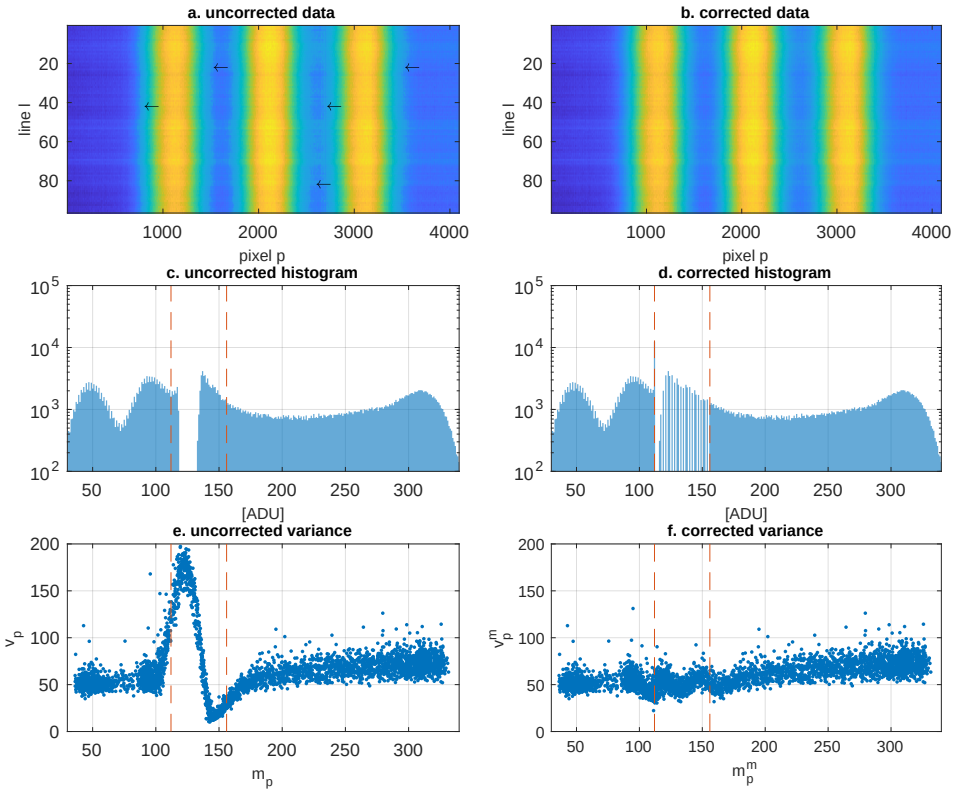


Figure 2.5: Measurement and correction of missing symbols.
(a) Scan of a defocused resolution target in false colors. The sensor is operated in 1xTDIx1 mode, an Analog Gain of 10 and sensorlet 64 is used. Displayed is color channel 1. The effect of the missing symbols is visible as sharp boundary between the dark background and the bright area's.
(b) Data after correcting for missing symbols. The boundary is not visible anymore.
(c-d) Fitted mapping function $M(X)$ and the corresponding inverse mapping function $M^{-1}(Y)$ for the missing symbols. Found was $X_1 = 112\,\text{ADU}$, $X_2 = 156\,\text{ADU}$ and $d = 19\,\text{ADU}$.
(e-f) Histogram for the raw- and the corrected data.
(g-h) Variance versus mean for the raw- and the corrected data.

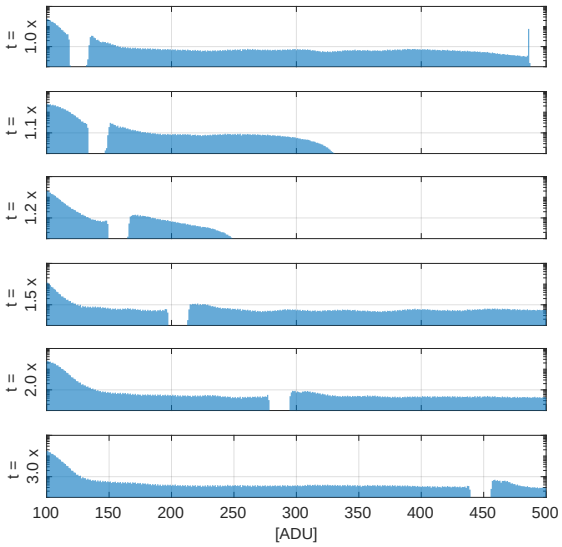


Figure 2.6: Missing symbols as a function of timing. Shown are the results for the first channel. The sensor was operated in 1xTDIx1 mode using sensorlet 64, an Analog Gain of 6 and an Analog Offset of 113. All timing parameters (LineTriggerPeriod, ChannelTriggerPeriod, ExposureTime), are increased with a factor ranging from 1 to 3 as indicated at the left of the graphs. The range of the missing symbols shifts proportionally to the right

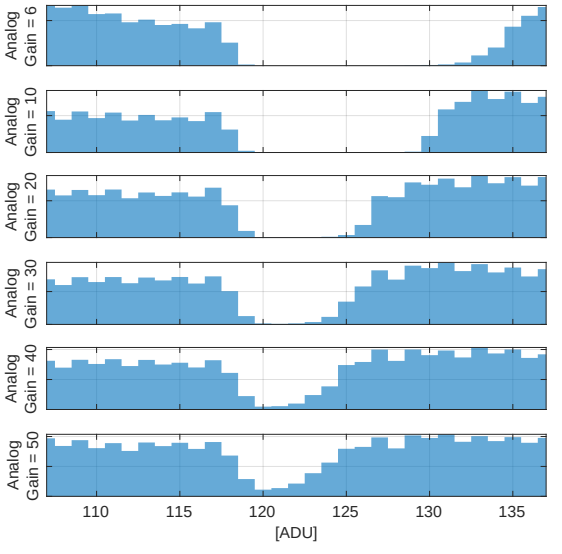


Figure 2.7: Missing symbols as a function of Analog Gain. The range of missing values becomes smaller for larger values of the Analog Gain.

dead time between the acquisitions was chosen to be about three times the exposure time. Because every exposed acquisition is preceded by a series of 8 unexposed acquisitions, we can make the approximation that, for any reasonable value of α, the expectation value of an exposed pixel is

$$E[X_1 - B_1] = Q_1 + \alpha^9 Q_{-8} \approx Q_1. \tag{2.16}$$

For the unexposed consecutive acquisition we than have

$$E[X_2 - B_2] = \alpha Q_1 = \alpha E[X_1 - B_1]. \tag{2.17}$$

By fitting the measured data against this linear model, a value for α can be obtained. As a demonstration of the image lag measurement, Fig. 2.8 shows the value of $x_{l,2000,1}$ with $l \in 1,4,\cdots$, i.e. the value of pixel 2000 of the first channel for the exposed lines. Also shown is the value of the consecutive acquisition $x_{l,2000,2}$. After about 2000 lines, the illumination pulse fully overlaps with the pixels integration period and the measured value no longer increases. Looking closely at the consecutive channel reveals that it also stops increasing. This proves the observed behavior is not caused by the illumination. Figure 2.8b shows the correlation between $x_{l,2000,1}$ and $x_{l,2000,2}$. A linear fit gives a slope of 6.7 %.

Image lag was measured for all pixels in the first channel by correlating x_{lp1} with x_{lp2}, where only channel 1 is exposed and $l = 1,4,\cdots$ corresponds to the exposed lines. This measurement was repeated for the second channel by illuminating it and correlating x_{lp2} with x_{lp3}, and for the third channel by illuminating it and correlating x_{lp3} with $x_{l+1,p1}$, i.e. the first channel of the next line. The results are given in Fig. 2.9. The image lag shows a small dependence on the pixel number. On average, the image lag was determined to be 6.7 %. Note that the found image lag is very similar for all channels, which shows that this is most probably an effect caused by the sensor. If this would have been caused by the readout electronics, it would be likely to see a dependence on the color channel.

In practice, image lag will make the red and green mix into the green and blue channel respectively while the blue channel mixes into the red channel of the next line. A mitigation might be found in linear unmixing of the channels.

2.4.5. Dark current

The impact of dark current is measured by acquiring data using an unexposed sensor for a range of exposure times. The values m_p and v_p are calculated for each scan. From Eq. (2.8) it follows that

$$\frac{\mathrm{dE}[X]}{\mathrm{d}\tau} = 0 \approx \frac{\mathrm{d}m_p}{\mathrm{d}\tau}, \tag{2.18}$$

$$\frac{\mathrm{dvar}[X]}{\mathrm{d}\tau} = \frac{I_d}{g_0} \approx \frac{\mathrm{d}v_p}{\mathrm{d}\tau}. \tag{2.19}$$

This implies that, due to the built-in dark current correction, m_p will not increase, but that v_p will increase due to dark current [25]. By linear regression of v_p against τ, a value of I_d can be obtained.

For our measurement, a range of exposure times between 42 and 211 µs (5 and 25× default) is used. Figure 2.10a shows the slope of the linear fit between m_p

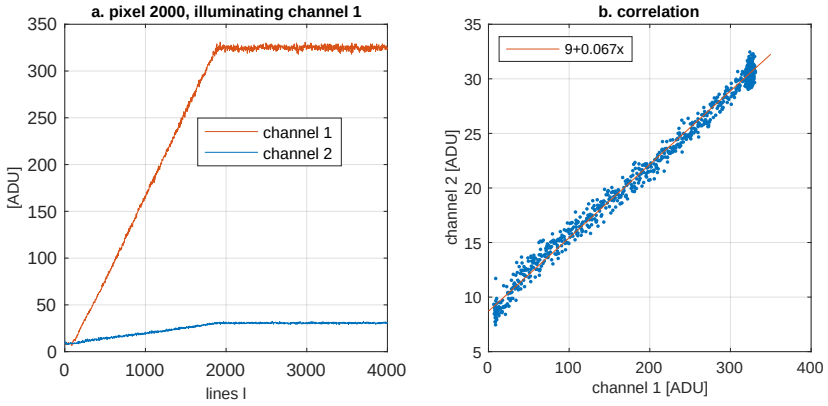


Figure 2.8: Demonstration of the image lag. The sensor was operated in 1xTDIx1 mode using an Analog Gain of 70 and an offset of 115. Sensorlet 64 was used. The exposure time was 42.2 µs (5× default) while the line trigger period was increased to have a dead time of about 160 µs between every acquisition. Only the first channel of one in three lines is illuminated. This graphs show only the values for $l = 1, 4, \cdots$, corresponding to the exposed lines.
(a) Value of pixel 2000 of the illuminated channel 1 (red curve) and the consecutive channel 2 (blue curve).
(b) The same data as scatter plot, showing the value of the consecutive channel versus the illuminated channel. Included is a linear fit indicating an image lag of 6.7 %.

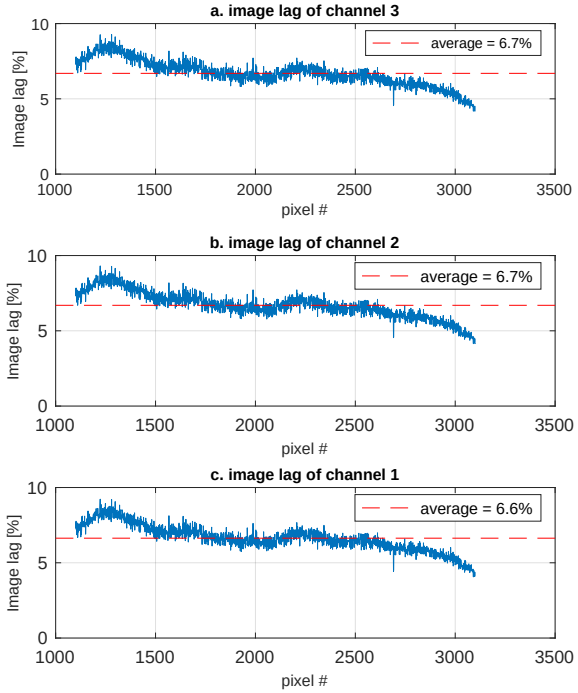


Figure 2.9: Image lag per pixel for every channel. Only pixels numbers 1100-3100 where included because of the limited size of the light spot used in this experiment. An average image lag of 6.7 % is found. The sensor operation mode is given in the caption of Fig. 2.8

and τ for all pixels. As anticipated, this value has a very low dependence of about 1 ADU/ms on average. Given that the maximum exposure time of the sensor is about 200 µs, this is negligible in all practical purposes. The increase of v_p as a function of exposure time is shown in Fig. 2.10b. An average of $1535\,\mathrm{e}^{-1}$/ms is found. A gain of $g_0 = 52.9\,\mathrm{e}^-$/ADU is used to convert variance to photoelectrons, see Section 2.4.8. To our current understanding, this number is unexpectedly high and might be caused by other mechanisms that are not yet understood. Regardless the underlying mechanism, the impact on the noise is the same.

The measured dark current implies that at the default exposure time (8.432 µs), the dark current will lead to 12.9 excited electrons which will contribute $3.6\,\mathrm{e}^-$ noise, which is a relatively small contribution. At the maximum exposure time of 210 µs (25× default), 324 electrons will be excited which will contribute $18\,\mathrm{e}^-$ to the sensor noise.

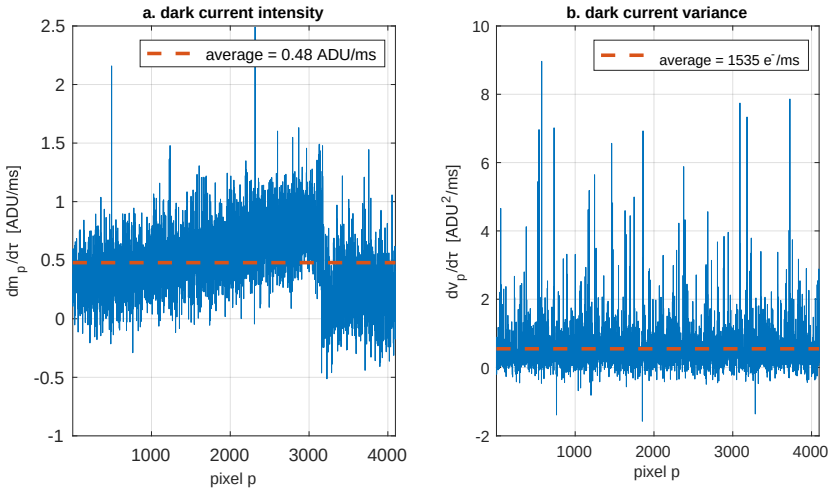


Figure 2.10: Effects of Dark current. Measured in 1xTDIx1 mode using sensorlet 64, an Analog Gain of 32 and an Analog Offset of 116
(a) Background intensity increase due to dark current.
(b) Dark current measured by the increase in variance as a function of exposure time.

2.4.6. Background

The sensor background was measured by acquiring a scan of an unexposed sensor. An illustration of the result is given in Fig. 2.11a. Now, by calculating $B(y) \approx m_p$, the static background is found. Figure 2.12 shows the static background for a range of Analog Gains and Analog Offsets. In Fig. 2.11b, the static profile is removed to give an impression of the fluctuating background. Although this background fluctuation is random, it does not show a dependence on the pixel. The fluctuating background profile can therefore be obtained from the dark pixels. Good results where obtained by averaging the center 24 black pixels of the left column:

$$B_f(t) \approx \mathrm{mean}_p[b_{pl}], p \in 4 \cdots 28. \tag{2.20}$$

This is used to correct for the fluctuating background profile. The last Fig. 2.11b shows the residual background after background removal. For background corrected values, we will use the notation with a b superscript

$$x^b_{lp} = x_{lp} - m_p - \text{mean}_{p'}[b_{p'l}], p' \in 4,5\cdots 28, \tag{2.21}$$

$$m^b_p = \text{mean}_l \left[x^b_{lp} \right], \tag{2.22}$$

$$v^b_p = \text{var}_l \left[x^b_{lp} \right] \tag{2.23}$$

where m_p is obtained from a scan of an unexposed sensor using the same settings and b_{pl} from the same scan as x_{pl}. Both the static and the fluctuating background do no impact image quality when they are corrected for. This is possible provided that the sensor operates linearly. This means, for example, that having missing symbols in the relevant range for the black pixels should be prevented.

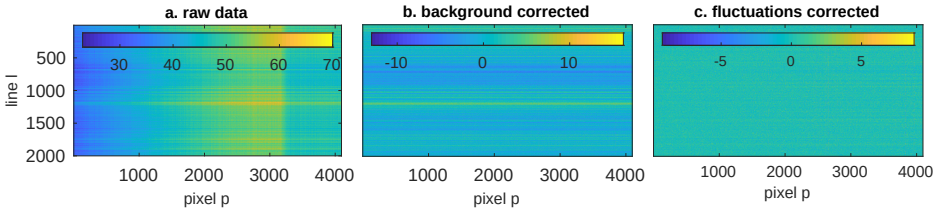


Figure 2.11: Illustration of the background profile. Measured by acquiring 2000 lines in 1xTDIx1 mode using sensorlet 64, while the sensor was not exposed.
(a) Raw data showing an unexposed image at Analog Gain = 30 and offset = 115
(b) Image with static background profile removed.
(c) Image with both static and fluctuating background profile removed.

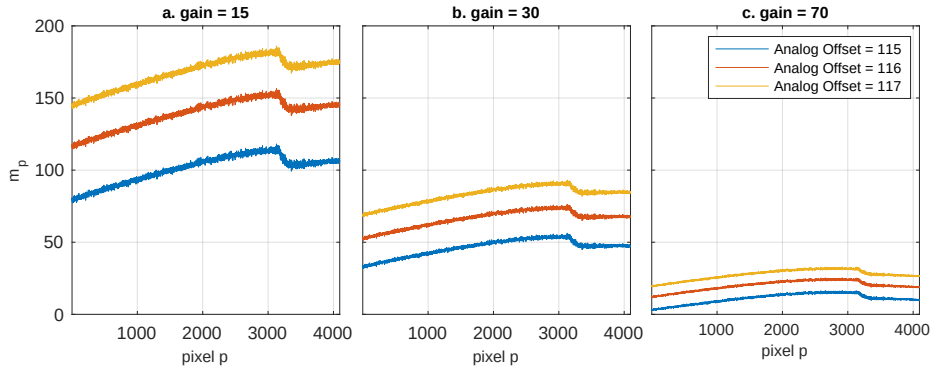


Figure 2.12: Background profile at an Analog Gain of 15 (a), 30 (b), 70 (c) for an Analog Offset of 115 (blue), 116 (orange) and 117 (blue).

2.4.7. Gain fluctuations

Fluctuations in the read-out amplifier are investigated by measuring a series of frames of a static defocused resolution target. The measurements are cor-

rected for the background profile. The resulting image is shown in Fig. 2.13a. Three 'bins' of pixels where selected based on their average intensity: a bin of 0-50 ADU,representing the darkest pixels, 100-150 ADU,where the fluctuations become visible, and the brightest pixels 400-450 ADU, see Fig. 2.13b. For each bin, the noise spectrum s_n was calculated using

$$s_n = \text{mean}_p \left[\mathscr{F}_l \left[x_{lp}^b \right] \right], \text{with } p \text{ such that } \begin{cases} 0 \le m_p^b \le 50 & \text{for } n = 1 \\ 100 \le m_p^b \le 150 & \text{for } n = 2 \\ 400 \le m_p^b \le 450 & \text{for } n = 3 \end{cases} \tag{2.24}$$

where n is the bin number. The result is shown Fig. 2.13c. The noise due to gain fluctuations clearly stands out on top of the flat spectrum of quantization-, shot- and read noise.

The gain fluctuations have an impact on pixel variance v_p^b. This is investigated using a measurement of a static defocused resolution target. We make use of the fact that the gain fluctuations are found to affect only a narrow range of frequencies. We therefore apply a band-stop filter to the signal in order to suppress the affected frequencies without distorting the remaining frequency content of the signal. The variance due to white noise sources will then not be reduced. This implies that the transfer function of this digital filter H is defined by

$$H(0) = 1, \tag{2.25}$$

$$H(f) = 0, f \in \text{affected frequencies}, \tag{2.26}$$

$$\langle H(f)^2 \rangle_f = 1. \tag{2.27}$$

The first requirement ensures that the mean over the lines will be unaffected. The filtered variance is denoted by a superscript f and is given by

$$v_p^f = \text{var}_l \left[H \otimes_l x_{lp}^b \right], \tag{2.28}$$

where $\otimes_l$ denotes convolution over the l dimension. In Fig. 2.14, a scatter plot is shown of the unfiltered mean m_p^b and variance v_p^b. Also the filtered variance v_p^f is plotted. The unfiltered curve deviates from the filtered values starting around 70 ADU. This also creates a discontinuity that is hardly present in the filtered data. This figure allows to quantify the impact of the gain fluctuations. At an intensity of 400 ADU, the variance is increased from about 8 to about 10 ADU^2. Expressing this in rms photoelectrons, the temporal noise leads to a $\sqrt{10/8} = 12\,\%$ increase of noise in the white parts of the image.

Further tests show that the gain fluctuations are invariant to the Analog Gain, invariant to the Analog Offset, independent of the chosen sensorlet, independent of the pixel position and appearing in all color channels. It is present in both 4xTDIx1 mode as well as in 1xTDIx1 mode. The distinct spectral peaks only show in the scan (i.e. time) direction. This effect appears to be unrelated to the fluctuating background as they have an uncorrelated spectrum. The position of the noise in the spectrum is highly dependent on the line rate, suggesting that this is a high frequency noise source which is aliased. As this effect is independent of the Analog Gain setting, it is independent of the absolute number of photoelectrons, which

suggests that this is not reset noise, but rather electronic interference. This noise will significantly impact image quality, especially in the white areas of the image, where it leads to a striping artifact oriented perpendicular to the scan direction.

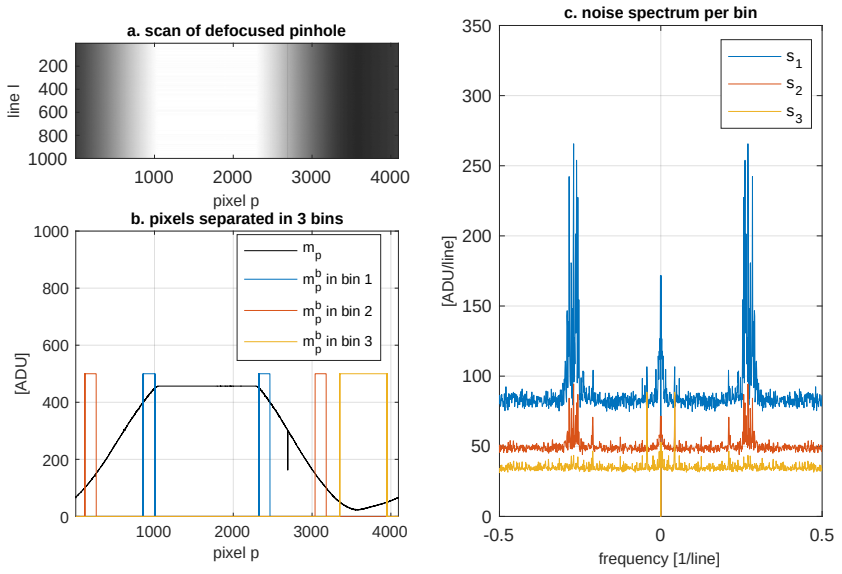


Figure 2.13: Demonstration of the temporal noise.
(a) A scan is made of a defocused pinhole (1xTDIx1, gain=32, exposure time = 101.2 µs, 12× default). Fluctuations are removed.
(b) The pixels are separated in 3 bins, with low, medium and high intensity.
(c) The average noise spectrum along the time axis is calculated for each bin.

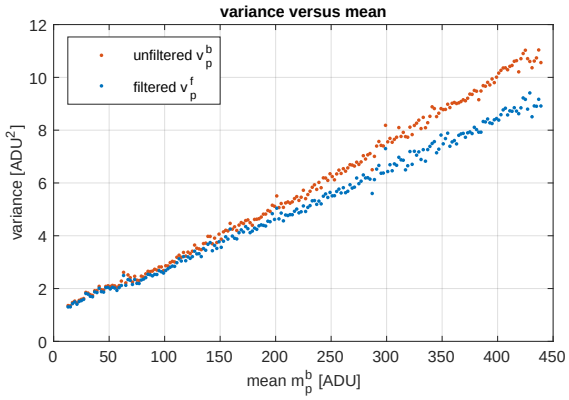


Figure 2.14: Impact of the temporal noise on the sensor calibration. Shown are the Variance versus Mean values and the linear regression for the unfiltered (red) and filtered (blue) image data. Sensor was run in 1xTDIx1 mode, with an Analog Gain of 32, and an Analog Offset of 116 and an exposure time of 101.2 µs (12× default).

2.4.8. Sensor gain

An essential step in the experimental evaluation of the sensor is the determination of the sensor gain and noise. This is done using the photon transfer method [19,

20], which exploits the linear relation of the shot noise induced variance with the number of photoelectrons. Provided that the sensor is linear, and, except for shot noise, only signal independent additive noise sources are present with a standard deviation σ, the total variance is given by

$$g^2 \mathrm{var}[X] = g\mathrm{E}[X-B] + \sigma^2 \tag{2.29}$$

The sensor gain g_0 and standard deviation of additive noise σ are measured from a scan of a static defocused resolution target. The data is background corrected and gain fluctuations are filtered out using the band-stop filter H defined by Eq. (2.27). To prevent image lag to affect the estimation of the variance, the data of only one channel is used. This reduces the correlation between the measured values due to image lag to α^3 which is negligible in practice. Therefore, we can approximate

$$\mathrm{E}[X-B] \approx m_p^b, \tag{2.30}$$

$$\mathrm{var}[X] \approx v_p^f. \tag{2.31}$$

A linear fit is subsequently performed between v_p^f and m_p^b to obtain values for g and σ.

The sensor gain was measured for a wide range of Analog Gain settings. The results are shown in Fig. 2.15 and are summarized in Fig. 2.16. Figure 2.16a shows that the gain depends linearly on the Analog Gain setting, which is the expected behavior. The sensor is hardly usable below an Analog Gain setting of about 5 because the Analog Offset was either too high (only clipping) or too low (underexposure). This means the minimum gain of the sensor is around 10 e^-/ADU. Figure 2.16b shows the system noise σ as a function of the Analog Gain setting. By correcting the total noise for quantization noise and dark current noise, the contribution of read noise can be calculated using $\sigma_0^2 = \sigma^2 - I_d\tau - g_0^2/12$. The measured read-noise shows a rather irregular shape, but has its minimum around an Analog Gain setting of 20.

The sensor gain was also measured with both engines running in 4×TDI×1 mode for an Analog Gain of 32. Figure 2.17 shows a substantial difference in gain between the ADCs. This is most probably caused by a difference in the on-board amplifiers and is not a result of a software configuration. At the time of writing, setting a different Analog Gain per ADC is not supported by the driver software.

2.4.9. Linearity of response

The linearity of the sensor is verified by a scan that is made using a flat and constant illumination for a range of exposure times. The results are displayed in Fig. 2.18a. For every pixel a linear fit of the background corrected averaged value m_p^b with the exposure time τ was done, as is illustrated in Fig. 2.18b. The RMS error of the linear fit was calculated for each pixel and is given in Fig. 2.18c. On average an error of only 0.7 ADU was found. Also the 95-percentile was calculated, which remains below 1.4 ADU for all pixels. There is a small but significant relation between the error value and the position on the sensor which is not fully understood. For practical purposes, however, the deviation of linearity is low and the sensor can be regarded linear over the full dynamic range.

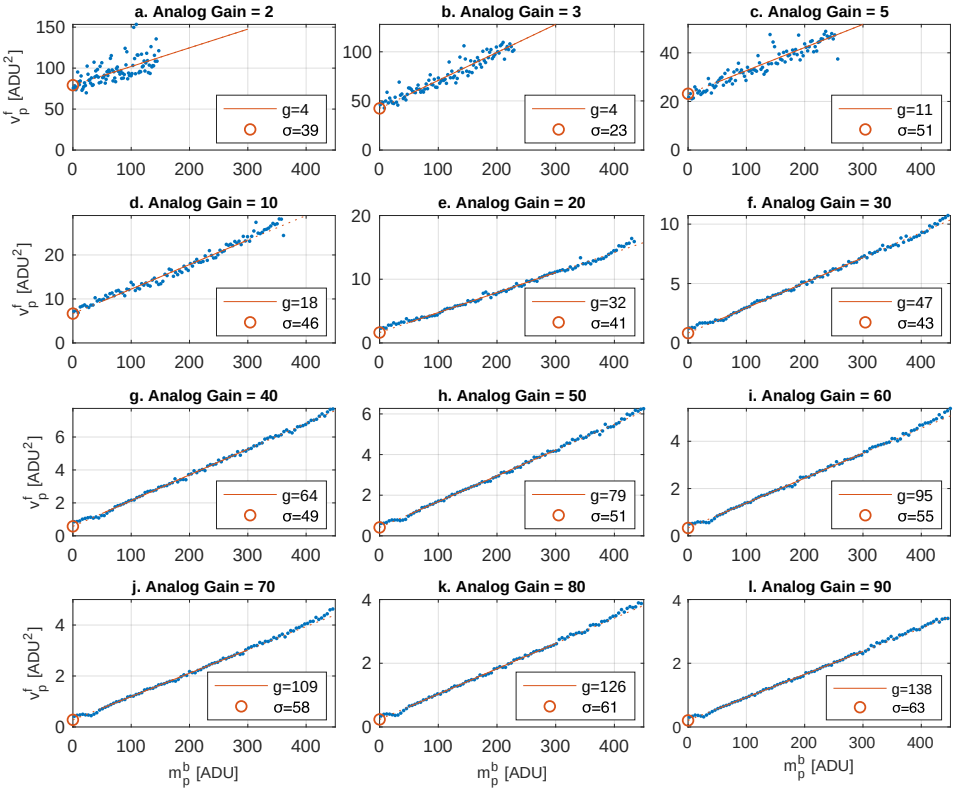


Figure 2.15: Sensor calibration for a range of Analog Gain settings. The obtained read noise σ is given in e^- and the gain g in e^-/ADU. The sensor was operated in 1xTDIx1 mode using sensorlet 64. Background and fluctuation correction was applied, and temporal noise was filtered out. An exposure time of 84.3 µs (10× default) was used.

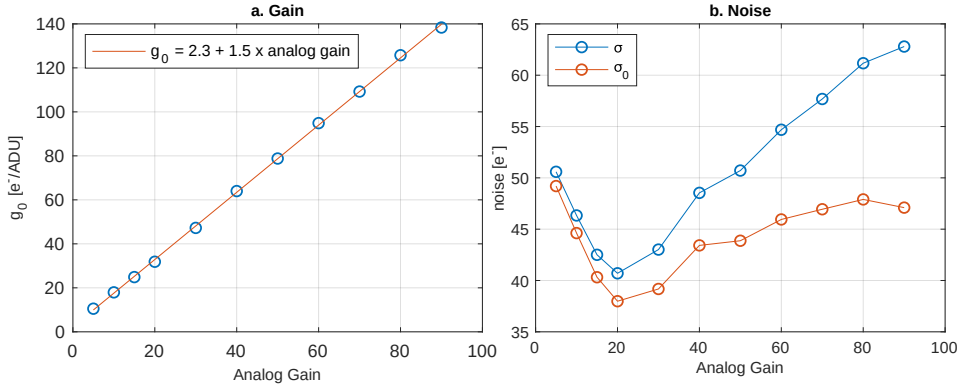


Figure 2.16: Summary of the sensor calibration results. Full measurement data is shown in Fig. 2.15.
(a) Sensor gain g_0 as a function of the Analog Gain setting including a linear fit.
(b) Total system noise σ as a function of Analog Gain. Also shown is the contribution of read-noise σ_0.

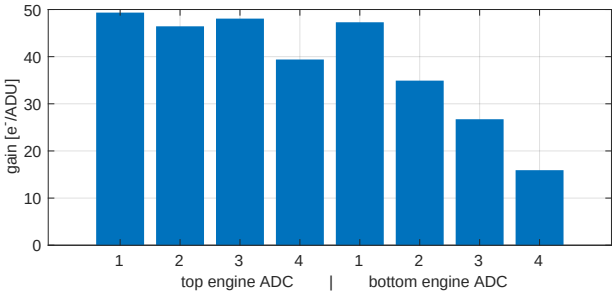


Figure 2.17: Gain with both engines in 4×TDI×1 mode with an Analog Gain of 32, an Analog Offset of 116 for the top engine and 124 for the bottom engine, and an exposure time of 101.2 µs (12× default). A filter for rejection of the temporal noise was applied. Bar 1-4 correspond to the ADCs of the top engine. Bar 4-8 correspond to the ADCs of the bottom engine.

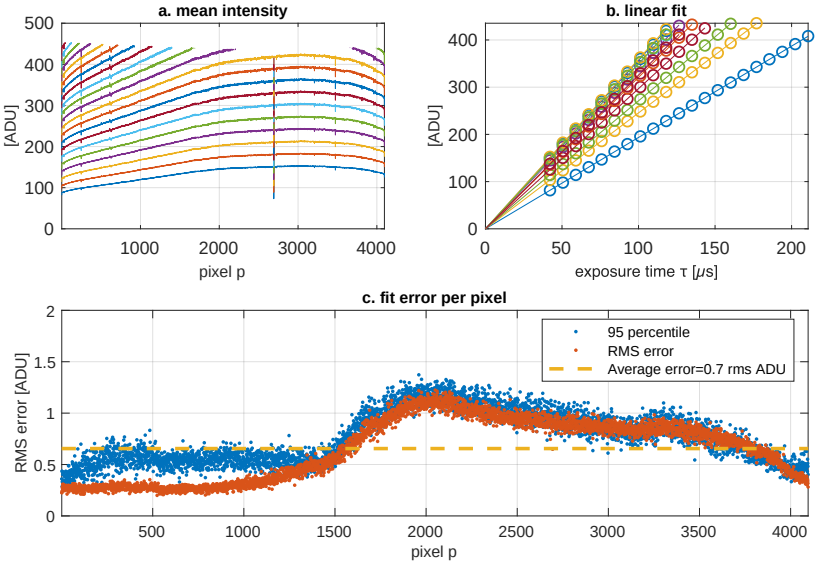


Figure 2.18: Linearity of the sensor response.
(a) Raw data showing the mean intensity of a scan of 1000 lines in 1xTDIx1 mode, Analog Gain=32, Analog Offset=116 using sensorlet 64 for a range of exposure times between 42 µs and 210 µs (5× and 25× default). Fluctuation and background correction was applied. Values above 450 ADU where disregarded to prevent clipping effects.
(b) Intensity as a function of exposure time for a subset of pixels, together with a linear fit.
(c) The RMS error and 95-percentile of the deviation of the linear fit for every pixel.

2.4.10. Lateral crosstalk

The sensor was tested for lateral crosstalk by imaging a slightly defocused resolution target onto the right side of the sensor, while the left side of the sensor was illuminated without obstruction. The values of the unobstructed pixels where correlated to the modulated pixels to assess the correlation. A lack of correlation indicates the absence of (long distance) lateral crosstalk. Figure 2.19 shows the result for the lateral crosstalk measurement. The values of the pixels 100, 200, 300, 500, 1000, and 1500 are plotted as a function of average value of pixels 3000 to 3500, see Fig. 2.19b. It can be noted that there is no relation between the value of these pixels and the intensity of the modulated side of the sensor. This indicates that there is no lateral crosstalk present in this image.

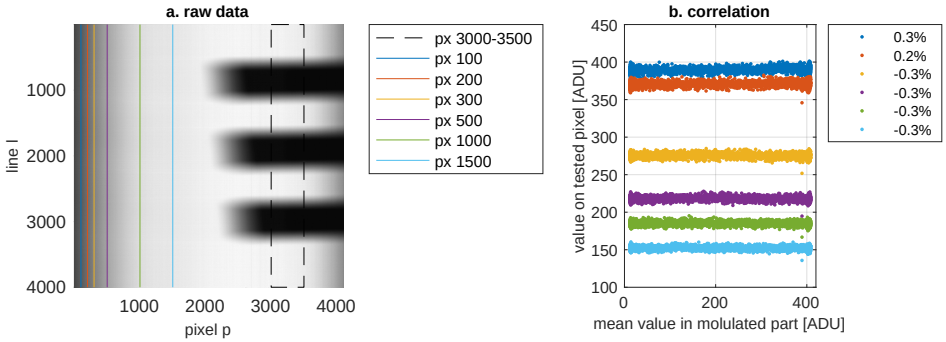


Figure 2.19: Test for lateral crosstalk. Sensor was operated in the 1×TDI×1 mode, an Analog Gain of 32 and an Analog Offset of 116. Sensorlet 64 was used.
(a) Raw data showing a slightly defocused resolution target imaged at the right of the sensor.
(b) Values of pixels 100, 200, 300, 500, 1000, and 1500 plotted as a function of average value of pixels 3000 to 3500. The legend shows the calculated crosstalk for the indicated pixels

2.5. Discussion

2.5.1. Implications for achievable SNR

The measured gain, dark current and read-noise can be used to determine the SNR as a function of the available photo-electron current, using Eq. (2.9). As an illustration of the different sources of noise, Fig. 2.20 shows the noise contributions as a function of the photo-current for the default Analog Gain and exposure time. The smallest noise contribution is from the dark current shot noise, which is about two orders of magnitude smaller than the read noise. Even for the longest possible exposure time on the system, the dark current shot noise will remain below the quantization noise. This implies that increasing the exposure time will in practice hardly increase the system noise. Also, the quantization noise of the system is much lower than the read noise for the given Analog Gain.

Now that all noise contributions are known, we can calculate the optimal Analog Gain setting given the available photo-electron current. The optimal setting is found by maximizing the SNR as a function of the Analog Gain, with the boundary condition that the measured value does not exceed 450 ADU. The results are displayed in Fig. 2.21a, for both the minimum exposure time (8.432 µs) and the maximum exposure time (210.8 µs). We found the read noise to have a minimum

at an Analog Gain setting of 20 (see Fig. 2.16). Due to the dominant role of the read noise (see Fig. 2.20), it is never advantageous to decrease the Analog Gain to reduce the quantization noise. This means that, although lower gains can be achieved, it will never be advantageous to have a gain lower than 32 e^-/ADU. For higher photo-electron currents, keeping the Analog Gain at 20 would lead to clipping and higher values must then be chosen.

Figure 2.21b shows the corresponding SNR values, for both the minimum exposure time (8.432 µs) and the maximum exposure time (210.8 µs). The results can be compared to the dashed line that indicates the theoretical limit of a shot noise limited SNR. In both curves, three regions can be discriminated. For very low photo-electron currents, the read noise of the system is dominant. The SNR increases therefore proportional with the signal. An SNR of 10 is obtained starting at photo-electron currents larger than 2 e^-/µs when the system is used at maximum exposure time. The second region is there where shot noise takes over as the dominant noise source, which is at 10 e^-/µs for the maximum exposure time and at 3×10^2 e^-/µs for the minimum exposure time. This point indicates the start of the region in which the sensor performs well. The SNR now increases proportional to the square of the photo-electron current. The third region starts where the Analog Gain has to be increased to prevent clipping. This is at about 70 e^-/µs for the maximum exposure time and 2×10^3 e^-/µs for the minimum exposure time. Here, the system is very close to shot noise limited. Increasing the photo-electron current beyond this point still increases the SNR, but the need for higher Analog Gain values will introduce some extra noise. For these very high photo-electron currents, SNR values up to 200 can be achieved.

The mentioned results are calculated based on the 1×TDI×1 mode, but can be easily be extrapolated to 1×TDI×4 mode by multiplying the resulting SNR values by a factor 2. Our noise model predicts an SNR of 146 when the full dynamic range of the sensor is used at the default exposure time and Analog Gain. When combined with 4 TDI lines, an SNR of 292 is expected.

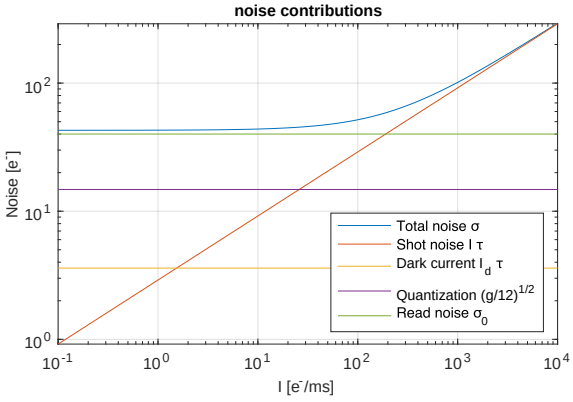


Figure 2.20: Calculated contributions of different noise sources as a function of the photo-electron flux. Assuming an Analog Gain of 32 (default) and an exposure time of 8.432 µs (default).

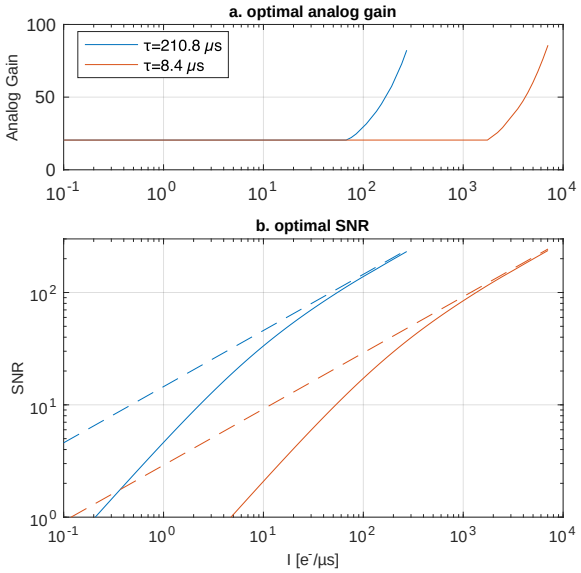


Figure 2.21: Calculation of the optimal sensor configuration as a function of the photo-electron flux for the minimum (blue) and maximum (red) exposure time.
(a) The calculated optimal Analog Gain.
(b) The corresponding SNR. The dashed line indicates the theoretical limit of a shot noise limited SNR.

2.5.2. Consequences for fluorescence scanning

We investigated the possibilities and limits of extending a WSI system to fluorescence imaging by using an LED based epi-illumination in recent research [23]. The results were obtained with the same setup as used in the current study. We found the photon yield to be between 0.2 and 1 $e^-/\mu m^2/\mu s$ depending on the sample and the fluorophore. Given the sensors target pixel size in object space of $(0.25\,\mu m)^2$, this corresponds to 0.01 to 0.06 $e^-/\mu s$. The sensor will not be able to capture meaningful images with such low photon yield, given the SNR (see Fig. 2.21). Improvements can be realized by combining the data of multiple sensorlets and by increasing the exposure time beyond current limitations. For example, by adding the signal of two sensorlets, for the least performing fluorophore considered in Ref. [23] it would require an exposure time of at least 2.6 ms to achieve an SNR larger then 1 and about 200 ms to have an SNR higher then 10. This is more than 20.000 times longer than the default exposure time. Although longer exposure times will increase the SNR, the system will not become shot noise limited, because the dark current is more than an order of magnitude larger than the signal. This indicates that the system is unsuitable for fluorescence imaging using an LED based illumination. Preliminary results with a laser line illumination source for fluorescence imaging suggests that the photon yield can be improved by two orders of magnitude. In that case, an SNR of 10 can be achieved using an exposure time of about 500 µs, on the order of the current maximum. This results in scan times that will be about 60 times slower than the default scan rate, which might be acceptable in practice. The read noise and dark current shot noise will be on the same order of magnitude as the shot noise.

2.5.3. Conclusions and next steps

The proposed slide scanner architecture with the custom CMOS multi-line sensor enables simultaneous auto-focus and image acquisition with 4-row TDI. Our system characterization shows that images with a very high SNR of 292 can be acquired, provided a sufficiently high photon-flux can be realized. This mean that 8-bit images will have a noise level less than 1 least significant bit (LSB). At the default settings, the gain is measured to be 50 e^-/ADU. The optimal gain in the limit of low photon currents is 32 e^-/ADU.

In our experimental characterization, however, some issues were found. At the high line rates, the sensor shows missing symbols, leading to non-linearities in the read-out. The resulting discontinuity has a size of about 200 e^-. We show that this effect can partly be corrected. The proposed method is computationally simple and therefore suitable for on-board correction. The validity of the model still has to be verified for higher Analog Gain settings. There might be room for improvement by considering a more complex model for the missing symbols, but this is possibly at the expense of an increase in computational complexity. A work around is to reduce the line rate to less than one fourth of the default, now at the expense of scanning speed.

The sensor shows high frequency fluctuations in the gain. This effect has a multiplicative nature and shows distinct spectral peaks. The location of these peaks depends strongly on the chosen line rate, suggesting this is a high frequency effect that is aliased multiple times. Numerical correction of this effect may be possible if a full description of the gain fluctuations $g_f(t,Q)$, introduced in Eq. (2.6), can be provided. Our measurements suggest that $g_f(Q,t)$ can be written as the product of a periodic function which we approximate by a single harmonic and a function describing the increase with intensity $C(Q)$:

$$g_f(t,Q) = C(Q)\cos(\omega t + \phi). \tag{2.32}$$

This model reflects that the fluctuations do not depend on the lateral pixel position and that the temporal part does not depend on Q. In order to explain why the gain fluctuations predominantly affect pixels with higher intensity, the function $C(Q)$ must be non-linear. A possible model is a step function

$$C(Q) = \begin{cases} 0 & \text{for } Q < Q_0 \\ C_0 & \text{for } Q > Q_0 \end{cases}, \tag{2.33}$$

with C_0 the amplitude and Q_0 a threshold value. This model suggests that only the phase ϕ remains to be determined for every measurement before it becomes possible to do a straightforward correction by a multiplicative factor depending only on line and intensity.

Some additional issues where found apart from the aforementioned two major issues. (1) The sensor shows a fluctuating background profile which can be completely compensated by using the black pixels, provided that the sensor behaves linearly. (2) The measured gain deviates across the ADCs, which may have an impact on the SNR. (3) We determined the image lag to be 6.7 % in a bright to dark transition. This can be solved by linear unmixing in a straightforward way. (4) We measured a rather high dark current of 1.5×10^3 e^-/ms. However, the attendant

noise does not have a significant impact on the overall SNR in practice. (5) We found the sensor to be linear with the exception of the earlier mentioned effects. (6) Finally, no lateral crosstalk between pixels in a row was found.

2.6. Acknowledgements

We acknowledge Bas Hulsken, Rik Kneepkens, Jelte Peter Vink, Mathijs Rem, Gert-Jan van der Braak and Aslihan Arslan Carisey at Philips Digital & Computational Pathology for support and feedback and for providing the multi-line image sensor for testing. NWO-TTW is thanked for financial support.

References

[1] J. Bury and J. Griffin. *Digital pathology*. Ed. by S. K. Suvarna, C. Layton, and J. D. Bancroft. 8th ed. Elsevier Health Sciences, 2018, p. 476. ISBN: 978-0-7020-6864-5 (cit. on p. 24).

[2] S. Al-Janabi, A. Huisman, and P. J. Van Diest. "Digital pathology: current status and future perspectives". In: *Histopathology* 61.1 (2012), pp. 1–9. DOI: 10.1111/j.1365-2559.2011.03814.x (cit. on p. 24).

[3] L. Pantanowitz, N. Farahani, and A. Parwani. "Whole slide imaging in pathology: advantages, limitations, and emerging perspectives". In: *Pathology and Laboratory Medicine International* 7 (2015), p. 23. DOI: 10.2147/PLMI.S59826 (cit. on p. 24).

[4] M. D. Zarella et al. "A Practical Guide to Whole Slide Imaging: A White Paper From the Digital Pathology Association". In: *Archives of Pathology & Laboratory Medicine* 143.2 (2019), pp. 222–234. DOI: 10.5858/arpa.2018-0343-RA (cit. on p. 24).

[5] C. Higgins. "Applications and challenges of digital pathology and whole slide imaging". In: *Biotechnic & Histochemistry* 90.5 (2015), pp. 341–347. DOI: 10.3109/10520295.2015.1044566 (cit. on p. 24).

[6] M. Montalto, R. Filkins, and R. McKay. "Autofocus methods of whole slide imaging systems and the introduction of a second-generation independent dual sensor scanning method". In: *Journal of Pathology Informatics* 2.1 (2011), p. 44. DOI: 10.4103/2153-3539.86282 (cit. on p. 24).

[7] N. Roberts et al. "Toward Routine Use of 3D Histopathology as a Research Tool". In: *The American Journal of Pathology* 180.5 (2012), pp. 1835–1842. DOI: 10.1016/j.ajpath.2012.01.033 (cit. on p. 24).

[8] Y. Xu et al. "A Method for 3D Histopathology Reconstruction Supporting Mouse Microvasculature Analysis". In: *PLOS ONE* 10.5 (2015). Ed. by A. H. Beck, e0126817. DOI: 10.1371/journal.pone.0126817 (cit. on p. 24).

[9] M. E. van Royen et al. "Three-dimensional microscopic analysis of clinical prostate specimens". In: *Histopathology* 69.6 (2016), pp. 985–992. DOI: 10.1111/his.13022 (cit. on p. 24).

[10] Y. Fan and A. P. Bradley. "A method for quantitative analysis of clump thickness in cervical cytology slides". In: *Micron* 80 (2016), pp. 73–82. DOI: 10.1016/j.micron.2015.09.002 (cit. on p. 24).

[11] B. Hulsken and S. Stallinga. "Sensor for microscopy". In: *US Patent No. 10,353,190 B2* (2019) (cit. on pp. 24, 25).

[12] B. Hulsken. "Scanning imaging system with a novel imaging sensor with gaps for electronic circuitry". In: *US Patent. No. 10,091,445 B2* (2018) (cit. on pp. 24, 25).

[13] B. Hulsken. "Method for simultaneous capture of image data at multiple depths of a sample". In: *US Patent No. 9,910,258 B2* (2018) (cit. on pp. 24, 25).

[14] S. Mendis et al. “CMOS active pixel image sensors for highly integrated imaging systems”. In: *IEEE Journal of Solid-State Circuits* 32.2 (1997), pp. 187–197. DOI: 10.1109/4.551910 (cit. on pp. 24, 29).

[15] A. El Gamal and H. Eltoukhy. “CMOS image sensors”. In: *IEEE Circuits and Devices Magazine* 21.3 (2005), pp. 6–20. DOI: 10.1109/MCD.2005.1438751 (cit. on p. 24).

[16] Hui Tian, B. Fowler, and A. Gamal. “Analysis of temporal noise in CMOS photodiode active pixel sensor”. In: *IEEE Journal of Solid-State Circuits* 36.1 (2001), pp. 92–101. DOI: 10.1109/4.896233 (cit. on pp. 24, 28, 29).

[17] E. Fossum. “CMOS image sensors: electronic camera-on-a-chip”. In: *IEEE Transactions on Electron Devices* 44.10 (1997), pp. 1689–1698. DOI: 10.1109/16.628824 (cit. on p. 24).

[18] K. Lidke et al. “The role of photon statistics in fluorescence anisotropy imaging”. In: *IEEE Transactions on Image Processing* 14.9 (2005), pp. 1237–1245. DOI: 10.1109/TIP.2005.852458 (cit. on p. 25).

[19] L. J. van Vliet, D. Sudar, and I. T. Young. *Digital fluorescence imaging using cooled charge-coupled device array cameras*. Ed. by J. E. Celis. 2nd ed. Vol. 3. Academic Press, 1998, pp. 109–120. ISBN: 978-0-12-164725-4 (cit. on pp. 25, 29, 41).

[20] J. R. Janesick, K. P. Klaasen, and T. Elliott. “Charge-Coupled-Device Charge-Collection Efficiency And The Photon-Transfer Technique”. In: *Optical Engineering* 26.10 (1987), p. 261072. DOI: 10.1117/12.7974183 (cit. on pp. 25, 41).

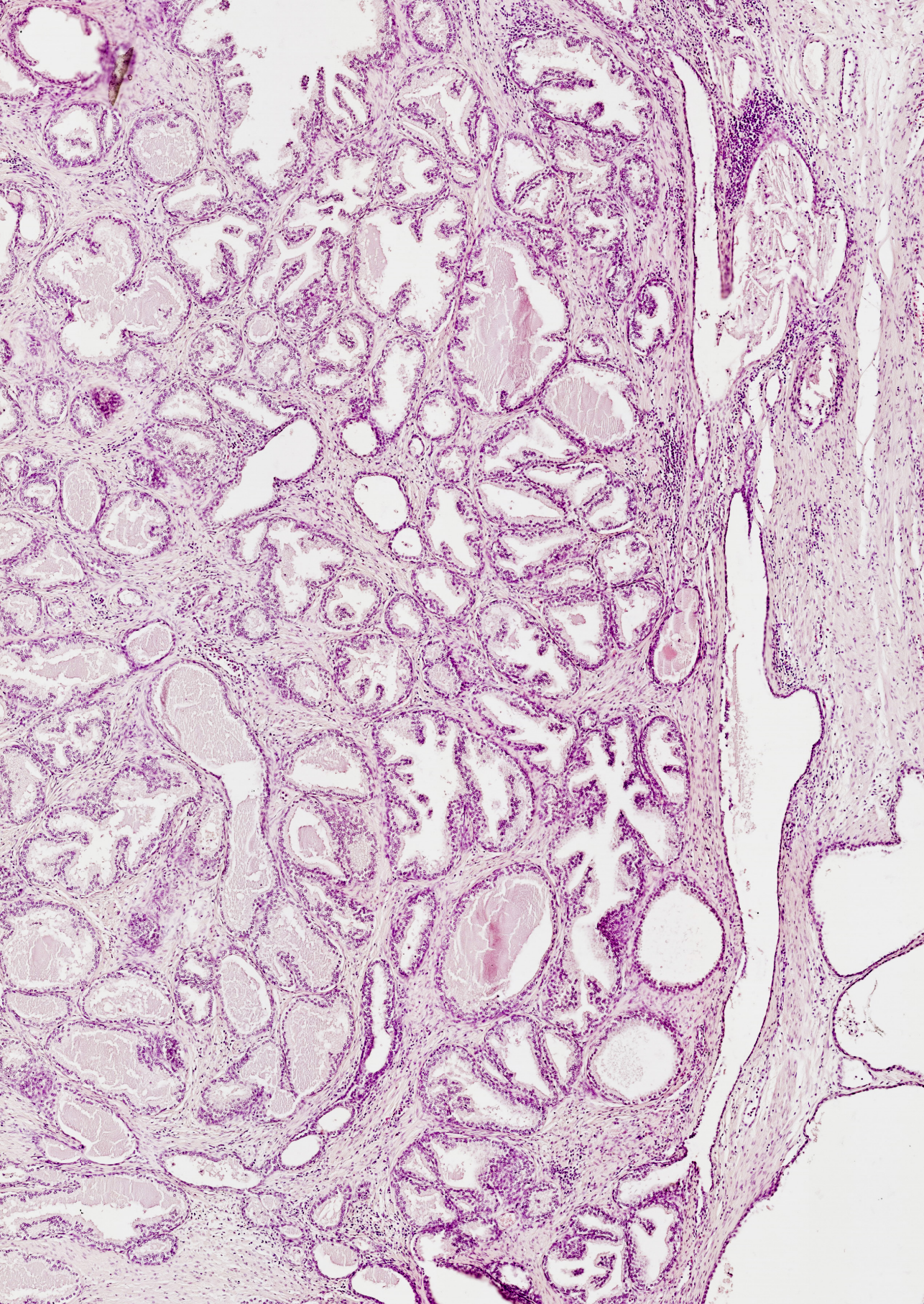
[21] J. C. Mullikin et al. “Methods for CCD camera characterization”. In: *Image Acquisition and Scientific Imaging Systems*. Vol. 2173. International Society for Optics and Photonics. 1994, pp. 73–84. DOI: 10.1117/12.175165 (cit. on p. 25).

[22] Q. Zhao et al. “Modulated electron-multiplied fluorescence lifetime imaging microscope: all-solid-state camera for fluorescence lifetime imaging”. In: *Journal of Biomedical Optics* 17.12 (2012), p. 126020. DOI: 10.1117/1.JBO.17.12.126020 (cit. on p. 25).

[23] L. van der Graaff et al. “Fluorescence imaging for whole slide scanning using LED-based color sequential illumination”. In: *Optics, Photonics, and Digital Technologies for Imaging Applications V*. Vol. 10679. International Society for Optics and Photonics. SPIE, 2018. ISBN: 9781510618848. DOI: 10.1117/12.2306776 (cit. on pp. 25, 30, 47).

[24] S. M. Shakeri et al. “Optical quality assessment of whole slide imaging systems for digital pathology”. In: *Optics Express* 23.2 (2015), pp. 1319–1336. DOI: 10.1364/OE.23.001319 (cit. on p. 25).

[25] H. Netten et al. “A fast scanner for fluorescence microscopy using a 2-D CCD and time delayed integration”. In: *Bioimaging* 2.4 (1994), pp. 184–192. DOI: 10.1002/1361-6374(199412)2:4%3C184::AID-BIO3%3E3.0.CO;2-M (cit. on pp. 26, 36).



3

Computational imaging modalities for multi-focal whole slide imaging systems

Whole slide imaging systems can generate full-colour image data of tissue slides efficiently, needed for digital pathology applications. This paper focuses on a scanner architecture that is based on a multi-line image sensor that is tilted with respect to the optical axis, such that every line of the sensor scans the tissue slide at a different focus level. This scanner platform is designed for imaging with continuous autofocus and inherent color registration at a throughput on the order of 400 MPx/s. Here, single-scan multi-focal whole slide imaging, enabled by this platform, is explored. In particular, two computational imaging modalities based on multi-focal image data are studied. First, 3D imaging of thick absorption stained slides (~ 60 µm) is demonstrated in combination with deconvolution to ameliorate the inherently weak contrast in thick-tissue imaging. Second, quantitative phase tomography is demonstrated on unstained tissue slides and on fluorescently stained slides, revealing morphological features complementary to features made visible with conventional absorption or fluorescence stains. For both computational approaches simplified algorithms are proposed, targeted for straightforward parallel processing implementation at ~ GPx/s throughputs.

This chapter has been published as: Leon van der Graaff, Geert J.L.H. van Leenders, Fanny Boyaval, and Sjoerd Stallinga in *Applied Optics* **59**, 20 (2020) [1].

3.1. Introduction

A higher resolution and a larger field of view has always been an important concern in the design of optical systems. The use of an objective lens with a higher numerical aperture (NA) provides a higher resolution but at the same time decreases the field of view (FOV) of the optical system. Whole Slide Imaging (WSI) systems provide a solution to this fundamental trade-off between resolution and FOV by scanning the tissue slide and acquiring a digital image with high resolution ($\sim 0.25\,\mu m$) with a truly unlimited FOV. WSI systems are the primary enabling technology for digital pathology, with applications in primary diagnosis [2–5], education [6–8] and research [9, 10].

Other methods are proposed to overcome the trade-off between resolution and FOV. Multi-spot scanning with array illuminators, here a large grid of high-NA spots is used for illumination, but a low NA with large FOV imaging path is used, have been considered [11–13]. In Fourier Ptychography [14, 15], a series of low resolution images of a large FOV acquired with different illumination beam angles are combined to produce a single high-resolution image of the entire FOV. Although these methods do overcome the resolution-FOV trade-off, they do not optimally use the space-bandwidth-time product of the optical system [16] (basically corresponding to the throughput in Mpx/s) because a fraction of the data stream is lost to overhead or carries redundant information. In addition, non-scanning approaches are still limited in FOV by the low magnification objective lenses used.

Currently, the majority of WSI systems provides high quality two-dimensional (2D) images of tissue slides, mostly for brightfield microscopy with white light illumination [17–19]. There are two main methods of scanning. The first is the 'step -and-stitch' method, in which the slide is moved step-wise and imaged using an area scanner. This is the most trivial extension to the traditional FOV-limited microscope but requires a potentially complex process of image-stitching in post-processing. The second is the 'push broom' or line scanning method, which combines simplicity and speed by scanning the slide with a line sensor at a constant velocity. The limited illumination etendue can be compensated by sensors with a Time Delay Integration (TDI) capability [20]. Current scanner systems have a throughput of several 100 Mpx/s, sufficient for imaging $\sim \mathrm{cm}^2$ tissue areas at sub-µm sampling density within ~ 1 min.

A major challenge in slide scanning is the need for tracking the topographic variations of the tissue layer which usually exceed the depth-of-focus of the scanner. To this end, a focus map of the whole slide can be made prior to scanning, where only a limited number of locations is used to save time [21]. An improvement is an autofocus system in which focus information is provided continuously, without mechanically changing focus. This information is then used for a closed-loop feedback system for maintaining optimum focus [22, 23]. Such an autofocus system can be realized using a second optical branch in the scanner for generating the focus information. Drawbacks of this approach are the increase in overall optical complexity and the required synchronization between the autofocus and imaging branches of the scanner.

Recently, a new WSI-system has been described in the patent literature by Philips [24–26] to overcome these drawbacks. The core of the approach lies in a new architecture for a multi-line image sensor. The sensor contains separate

'sensorlets', groups of adjacent pixel rows, which can be read-out independently. The sensor is tilted with respect to the optical axis resulting in a tilted object plane [27]. The readout of pixel information is done via two separate, simultaneously obtained, data streams, one obtained from a single sensorlet for providing the primary imaging information, the other obtained from multiple sensorlets for providing autofocus information. A full color whole-slide image with inherent color registration is acquired with a color-sequential illumination scheme based on fast switchable, high power LEDs [28]. In summary, the advantages of the proposed architecture are the elimination of the need for multiple cameras for imaging and autofocusing, the reduction of component costs and the simplification of manufacturing and maintenance.

An inherent capacity of the WSI platform is the possibility to make a multi-focal scan by readout of multiple sensorlets simultaneously [24–26]. This reduces scan time and has intrinsic axial registration of the scanned layers compared to acquiring a multi-layer image with multiple sequential 2D scans. The goal of this paper is to explore new imaging modalities of the scanner platform, based on this opportunity. Making 3D images with a single scan specifically leads to opportunities in imaging the 3D morphology of tissues and cells over large FOVs and in computational phase contrast imaging of unstained tissues.

Visualizing the 3D tissue morphology is needed in some cases in the field of histology before making a final diagnosis [17, 29]. Visualizing cells in 3D is generally always needed for diagnoses in the field of cytology [30]. The 3D imaging functionality of current WSI systems is mostly realized through sequential scans at different focus levels [31–33]. This approach leads to a considerable increase in scan time, and is sensitive to errors in registration of focal layers, which hamper its large-scale application. These disadvantages can be overcome by the proposed single scan 3D imaging mode.

Quantitative phase imaging forms a non-invasive and label-free imaging platform in cell biology and pathology [34]. Algorithms are available for phase retrieval from a through-focus image stack to obtain a 2D phase contrast image of a thin layer, usually based on solving the Transport of Intensity Equation (TIE) [35–38], as well as for a full 3D tomographic reconstruction of a thicker specimen [39–41]. Application of such computational phase contrast modalities based on multi-focal image stacks can broaden the application of WSI systems to unstained samples.

An important requirement for both modalities is the need for simple, scalable image processing methods, ultimately enabling real time image data processing. Key in achieving high throughput image processing is parallelization of the algorithm, such that it can be computed efficiently on a Graphics Processing Unit (GPU) or on dedicated hardware, e.g. a Field-Programmable Gate Array (FPGA). In this work the focus will therefore be on image processing algorithms that are compatible with large scale parallelization in a straightforward way. In particular, a deconvolution algorithm for improved sectioning in multi-focal volumetric imaging of thick slides and an algorithm for quantitative phase tomography are presented. In both a 'z-only' approach is proposed, in which the final deconvolved or phase image data is calculated independently for each lateral position (x, y), which leads to low memory, efficient, and scalable calculations.

This paper is structured as follows. First, the 3D scanner architecture is de-

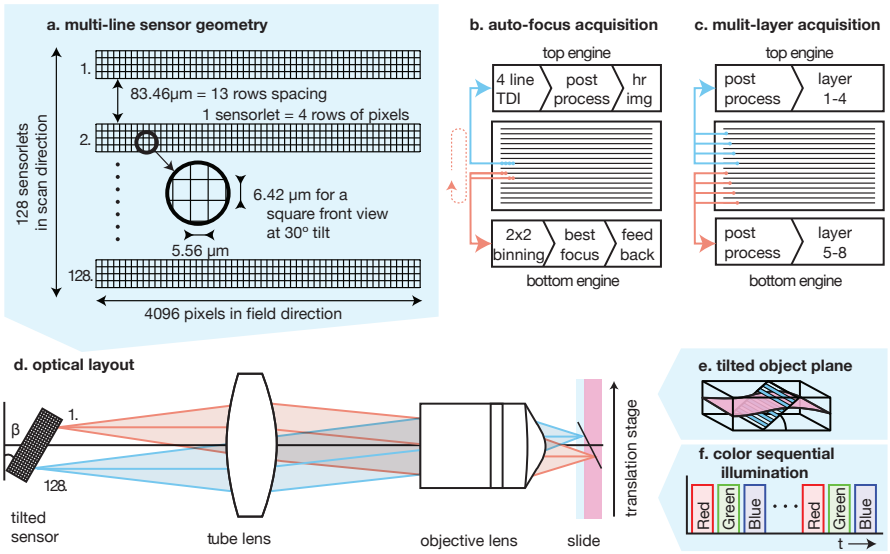


Figure 3.1: Schematic layout of the WSI system. (a) The pixel geometry of the novel multi-line CMOS image sensor. The sensor is equipped with two 'engines' that can independently read out the sensor. This can be used for (b) 2D scanning with continues autofocus or (c) 3D scanning. The optical layout (d) shows the tilted sensor that gives rise to a tilted object plane (e). A color sequential illumination strategy (f) is used for RGB imaging.

scribed including the novel image sensor. Next, multi-focal volumetric images of 60 µm thick tissue sections are shown. Then, the results of quantitative phase tomography of unstained tissue sections are presented. The paper is concluded with an evaluation of the findings and possible next steps for the highly modular WSI platform.

3.2. Scanner setup and image processing

3.2.1. 3D scanner architecture

A schematic outline of the 3D WSI platform is shown in Fig. 3.1. At the heart of the system lies a CMOS image sensor, of which the pixel geometry is outlined in Fig. 3.1a. The sensor consists of 128 sensorlets, groups of 4 adjacent pixel rows with a width of 4096 pixels. The sensorlets have a 13-row spacing. The pixel pitch is $\Delta_l = 6.42\,\mu\text{m}$ in the scan direction and $\Delta_p = 5.54\,\mu\text{m}$ in the orthogonal, or 'field' direction. The sensorlets have a 13 row spacing such that the sensorlet pitch $\Delta_s = (13+4)\Delta_l 115.56\,\mu\text{m}$. The sensor is tilted over an angle of $\beta = 30°$, so that each sensorlet scans the sample at a different depth, see Fig. 3.1e. This also results in a square projection of the pixels on the the plane of the tissue slide. The slide is illuminated by a color-sequential LED-based light source, equipped with a Köhler condenser [28], see Fig. 3.1f. The light source has three color channels for RGB imaging, with a typical wavelength of 657 nm for the red channel, 557 nm for the green channel and 465 nm for the blue channel.

Two Field Programmable Gate Array (FPGA) modules or 'engines' govern the readout of the sensor and provide the capability of on-chip processing of the image

data. Each engine has four Analog-to-Digital Converters (ADCs), which can be independently connected to a sensorlet of choice. The main purpose of having two engines is to use one for acquiring the image data, while the other engine is used to capture auto-focus data, as is illustrated in Fig. 3.1b. Optionally, the readout signal can be increased with a factor of four by application of TDI [20]. In this case, all four ADC's are connected to the four rows of a single sensorlet, and the FPGA applies digital TDI to create the primary, high resolution, high SNR, image information. Simultaneously, the other engine reads out a series of sensorlets sequentially, e.g. starting at sensorlet one, and increasing the sensorlet number with every readout line. Optionally, this is combined with 'binning' for higher signal levels, where the FPGA combines the data of two rows of a sensorlet to create a two times down sampled image with a four times higher signal level. The FPGA can calculate the optimal focus position from this data, and provides real time feedback to adjust the position of the objective lens. In this study we use an alternative readout mode in which data is acquired from 8 different sensorlets simultaneously, as is illustrated in Fig. 3.1c. For example, by reading out a single row of the sensorlets 1, 17,..., 127, an 8 layer multi-focal volume covering the full axial range can be obtained.

Images are acquired in a line-by-line or 'push broom' scanning fashion [42, 43]. The translation stages performs a linear translation with a velocity v_t while the sensor is triggered for data acquisition at equidistant positions with a step size Δ_t. We use $\Delta_t = 0.25\,\mu\text{m}$ and $v_t = 1\,\text{mm/s}$ resulting in a throughput of $4096{\cdot}3{\cdot}8{\cdot}v_t/\Delta_t \approx 393\,\text{MPx/s}$.

The objective lens and tube lens form a telecentric imaging system, see Fig. 3.1d. We use a Nikon 20× NA.75 Plan Apochromat VC objective lens ($F_{\text{obj}} = 10\,\text{mm}$) for high resolution imaging. Alternatively, a Nikon 10× NA.45 CFI Plan Apochromat λ objective lens ($F_{\text{obj}} = 20\,\text{mm}$) is used for a 2× wider scan lane and larger axial range, at the expense of lateral resolution. The tube lens has an effective back focal length of $F_{\text{tube}} = 222.4 \pm 2.2\,\text{mm}$ to match the sensor pixel pitch of $5.54\,\mu\text{m}$ with the target sampling density of $\Delta = 0.25\,\mu\text{m}$.

Scanning experiments were carried out on a prototype WSI system realized by Philips for research purposes.

3.2.2. Finite conjugate imaging and spherical aberration

A range of axial positions in object (sample) space is imaged onto the tilted detector by the imaging light path, comprising the objective lens and the tube lens. This imaging light path is optimized for imaging at a single depth inside the sample, typically directly after the cover slip. It follows that in this case we will suffer from spherical aberration. The sensitivity to spherical aberration can be analyzed along the lines of [44]. In that analysis, it is assumed that the objective lens and tube lens from an aplanatic telecentric imaging system.

We consider three degrees of freedom, see Fig. 3.2, namely z_1, the axial object position, z_2, the increase of the free working distance of the objective with respect to the nominal working distance, and z_3, the axial image position. These three distances are relative to the nominal aberration-free case. The total aberration function $W(\rho)$ is then given by:

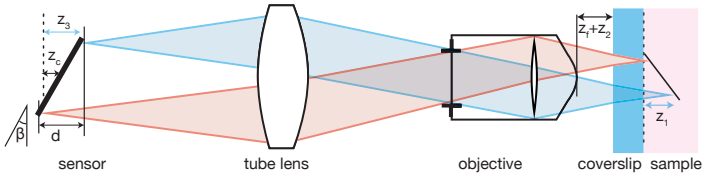


Figure 3.2: Illustration of the effect of the shifted axial position on focus and conjugate for two lines on the sensor (in blue and red). The light path runs from right to left so as to be consistent with previous figures. The sample consists of a coverslip (cyan) and the tissue of interest (pink) which are assumed to have a matched refractive index n. The dashed lines indicate the nominal object and image planes i.e. the focal plane of the tube lens and the surface below the coverslip. Three degrees of freedom are indicated: the axial object position z_1, the free working distance $z_f + z_2$ where z_f is the nominal free working distance, and the axial image position z_3. Tilting the sensor over an angle β makes it cover an axial range d in image space. Changing the axial position z_c of the sensor is used for minimizing the overall spherical aberration.

$$W(\rho) = z_1\sqrt{n^2 - \rho^2\mathrm{NA}^2} + z_2\sqrt{1-\rho^2\mathrm{NA}^2} - z_3\sqrt{1-\frac{\rho^2\mathrm{NA}^2}{M^2}}, \tag{3.1}$$

where ρ is the scaled radial pupil coordinate ($0 \le \rho \le 1$) such that the pupil is scaled to the unit circle, NA denotes the objective numerical aperture, $M = F_{tube}/F_{ob}$ is the lateral magnification, equal to the ratio of the tube focal length and the objective focal length, and n is the medium refractive index. This may be written in a compact form as:

$$W(\rho) = \sum_{j=1}^{3} z_j f_j(\rho), \tag{3.2}$$

$$f_1(\rho) = \sqrt{n^2 - \rho^2\mathrm{NA}^2}, \tag{3.3}$$

$$f_2(\rho) = \sqrt{1-\rho^2\mathrm{NA}^2}, \tag{3.4}$$

$$f_3(\rho) = \frac{\rho^2\mathrm{NA}^2}{2M^2}, \tag{3.5}$$

where a Taylor approximation is used for the third aberration term based on $M \gg 1$ and neglecting the piston term. The total amount of aberration is given by the root mean square (rms) value of the aberration function W_{rms}, which is given by:

$$W_{\mathrm{rms}}^2 = \sum_{j,l=1}^{3} g_{jl} z_j z_l, \tag{3.6}$$

with:

$$g_{jl} = \langle f_j f_l \rangle - \langle f_j \rangle \langle f_l \rangle. \tag{3.7}$$

Here the angular brackets indicate averaging over the pupil (i.e. integration over the unit circle with radial coordinate ρ). These averages can be evaluated analytically and are given in [44].

The degrees of freedom in our system can be reduced using two conditions. First, the object should be in focus. This is defined by the axial object position

z_1 for which the rms value of the aberration function is minimal, given the axial position of the sensor z_3 and the objective's working distance z_2. This implies that z_1 can be found by solving:

$$\frac{\partial W_{\text{rms}}}{\partial z_1} = 0, \tag{3.8}$$

which leads to:

$$z_1 = -\frac{g_{12}}{g_{11}} z_2 - \frac{g_{13}}{g_{11}} z_3. \tag{3.9}$$

This expression directly gives the axial magnification as:

$$M_{\text{ax}} = -\frac{g_{11}}{g_{13}} = \chi \frac{M^2}{n}, \tag{3.10}$$

where χ is a non-paraxial correction factor depending only on NA and n, which satisfies $\chi \to 1$ in the limit $\text{NA}/n \to 0$.

The second condition we impose is that the upper focal slice should be adjacent to the cover slip. The tilted image sensor spans a range of axial positions $z_c - d/2 \leq z_3 \leq z_c + d/2$ in image space, where d is the total axial range and z_c is the axial position of the center of the sensor. In object space, this corresponds to the axial range $0 \leq z_1 \leq d/M_{||}$. Now Eq. (3.9) implies that the working distance of the objective must be set such that:

$$z_2 = -\frac{g_{13}}{g_{12}} \left(z_c - \frac{d}{2} \right). \tag{3.11}$$

Using the expressions we derived for the object axial position and the working distance, see Eqs. (3.9) and (3.11), we can now write the rms value of the aberration function as a function of the axial image position, in which the axial position for the center of the sensor z_c remains as a free parameter. With some algebra it may be shown that:

$$W_{\text{rms}}^2 = \left(g_{33} - \frac{g_{13}^2}{g_{11}} \right) z_3^2 + 2 \left(\frac{g_{13}^2}{g_{11}} - \frac{g_{23} g_{13}}{g_{12}} \right) \left(z_c - \frac{d}{2} \right) z_3 + \left(\frac{g_{22} g_{13}^2}{g_{23}^2} - \frac{g_{13}^2}{g_{11}} \right) \left(z_c - \frac{d}{2} \right)^2. \tag{3.12}$$

Clearly, the aberration depends quadratically on the axial image position z_3, which is illustrated in Fig. 3.3. We can now choose z_c to minimize the overall aberration. This is achieved when the minimum of the parabola is at $z_3 = z_c$. This condition can be used to find an expression for z_c:

$$z_c = \frac{g_{13}^2/g_{11} - g_{23} g_{13}/g_{12}}{g_{33} - g_{23} g_{13}/g_{12}} \frac{d}{2}. \tag{3.13}$$

We will discuss the implications for the default $20 \times$ NA0.75 objective lens. The results for other optical configurations are given in Table 3.1. Using this objective, the lateral magnification is $M = 22.2$. The sensor has a pixel pitch of 5.56 µm, which gives a pixel pitch in object space equal to $p = 0.25\,\mu\text{m}$. Assuming a medium refractive index $n = 1.5$, it is found that $\chi = 1.07$ and the axial magnification is $M_{\text{ax}} = 352$. The image sensor size in the direction perpendicular to the line sensors is $127 \times 17 \times 6.42\,\mu\text{m} = 13.9\,\text{mm}$. With a tilt angle $\beta = 30°$ the axial range in image

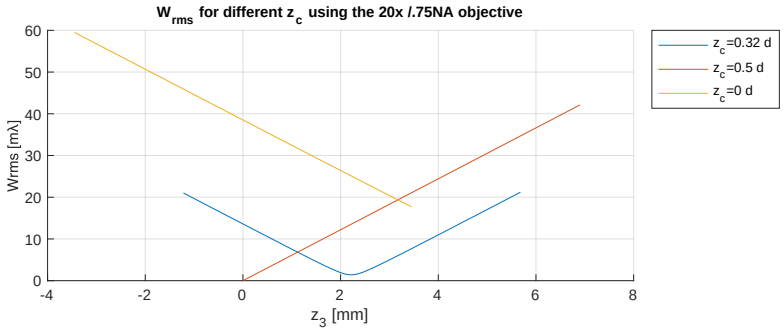


Figure 3.3: The total aberration W_{rms} as a function of the axial position in image (detector) space for different values of z_c: the optimal value ($z_c = 0.32d$), with the top of the sensor at the nominal plane ($z_c = d/2$), and with the center of the sensor at the nominal plane ($z_c = 0$). The curves are calculated using NA = 0.75 and λ = 500 nm.

space is $d = 6.9\,\text{mm}$ and the axial range in object space is $d/M_{\text{ax}} = 19.7\,\mu\text{m}$. The optimal axial position of the center of the sensor is $z_c = 0.32d$, giving an axial range in image space of $-0.18d \leq z_3 \leq 0.82d$. The reason for the significant asymmetry is the use of finite conjugate compensation of the spherical aberration arising from focusing into the sample below the cover slip.

Figure 3.3 shows the rms spherical aberration as a function of the axial position in image (detector) space. With the optimum choice for z_c, the rms spherical aberration for the center of the sensor is only $1.4\,\text{m}\lambda$ for a typical green emission wavelength of $\lambda = 500\,\text{nm}$, composed mainly of higher order spherical aberration, and $21\,\text{m}\lambda$ at the edge of the sensor. The spherical aberration varies in an asymmetrical way from the bottom of the sensor to the top of the sensor, in case the axial position of the middle of the sensor is not chosen optimally. For example, by choosing $z_c = 0$, the center of the sensor is at the nominal image plane. Now the rms spherical aberration ranges up to $60\,\text{m}\lambda$ for the worst-case edge of the sensor. When $z_c = d/2$ is chosen, the top row of the sensor is in the nominal image plane and has no aberration. However, the row at the other side of the sensor will then experience an rms aberration of $42\,\text{m}\lambda$.

Table 3.1: Calculated optical properties for optical configurations, with different magnification and axial range of the 3D scan. Shown are the used objective lens, the NA, the focal length F_{obj} of the objective lens, the lateral magnification M, the axial magnification M_{ax}, the non-paraxial correction factor χ, the lateral pixel pitch in object space Δ, the target sampling density in object space Δ_{ax}, the resulting axial range in object space d/M_{ax}, and the root mean square spherical aberrations W_{rms} at the center and edge of the sensor respectively. The values are based on $n = 1.5$, $\lambda = 500\,\text{nm}$ and $\beta = 30°$.

Obj.	NA	F_{obj} [mm]	M	M_{ax}	χ	Δ [µm]	Δ_{ax} [µm]	d/M_{ax} [µm]	W_{rms} [mλ]
20×	.75	10	22.2	352	1.07	0.25	0.15	19.7	1.4 - 21
10×	.45	20	11.1	84.1	1.02	0.50	0.65	82.4	0.2 - 10
4×	.20	50	5.55	13.2	1.00	1.25	2.65	512	0.0 - 2.4

3.2.3. Image post processing

In this work, no use is made of the on-chip processing capabilities of the sensor, but instead raw measurement data is written to a digital file for further processing in MATLAB. The four dimensional image data set $I[l,p,c,s]$ depends on a first index $l \in [1,N_l]$ for the 'scan' direction with size N_l equal to number of scanned lines, a second index $p \in [1,N_p]$ for the field direction, with $N_p = 4096$ pixels, a third index $c \in [1,N_c]$ representing the color channels, which has $N_c = 3$ for RGB imaging, and a fourth index $s \in [1,N_s]$, enumerating the $N_s = 8$ layers. Lines and channels are acquired line after line, and for every line N_c channels are acquired sequentially. The connotation of the term 'line' is therefore temporal in the current context and must not be confused with a physical row of pixels on the sensor.

Three pre-processing steps are taken for all data sets. First, the sensor offset is corrected, which consists of two terms. The sensor has a line-to-line varying offset $I_d[l,c,s]$ which is provided in the raw data of each scan. The second offset is a lateral offset $I_l[p,c,s]$, which must be measured only once for every particular configuration of exposure time, gain, and chosen sensorlets. The offset was measured by creating a 'dark' scan without illumination. The second pre-processing step is the registration that is required to compensate for the spatial offset in the scan direction of the different sensorlets used to scan the focus layers. Additionally, small registration errors arising from misalignment of the sensor tilt and rotation typically arise. This is compensated by translating the image data d_l pixels in the scan direction and d_p pixels in the field direction, given by:

$$d_l[s] = \left[\frac{N_t \Delta_s}{\Delta_l} \left(s - \frac{N_s+1}{2} \right) + \left(r - \frac{N_r+1}{2} \right) \right] D(1+\delta_l), \tag{3.14}$$

$$d_p[s] = \left[\frac{N_t \Delta_s}{\Delta_l} \left(s - \frac{N_s+1}{2} \right) + \left(r - \frac{N_r+1}{2} \right) \right] \delta_l, \tag{3.15}$$

where a constant sensorlet interval N_t is assumed e.g. $N_t = 4$ when sensorlets 50, 54, ..., 78 are used. D is the number of stage steps of size Δ_t it takes to translate the image to the next row on the sensor, given by $D = \Delta_l \cos(\beta)/M\Delta_t$. The stage step size Δ_t is chosen such that $D = 1$ for the 20× objective lens and $D = 2$ for the 10× objective lens. The value r is the row within the sensorlet at which the data is acquired, given by:

$$r[s] = \begin{cases} s, & s \leq 4, \\ s-4, & s \geq 5. \end{cases} \tag{3.16}$$

The overall displacement is minimized by correcting s and r for the number of used sensorlets $N_s = 8$ and the number of rows in a sensorlet $N_r = 4$. The residual alignments errors δ_l and δ_p are found by scanning a checkerboard resolution target and searching the value that optimizes the correlation between the layers. This is implemented by a coarse 5×5 grid search around zero with a step size 0.02, followed by a parabolic interpolation of the result. Finally, the image is scaled in the scan direction with a factor $\Delta_p/\Delta_t M$, to have isotropic sampling. The system is designed to have this factor equal to one for the 20× objective lens such that this scaling can be omitted in that particular case. The third pre-processing step is to correct for non uniformity of the illumination by applying flat fielding. The illumi-

nation intensity $I_f[l,c,s]$ is found by scanning a transparent area of a resolution target and average the result over the scan direction.

The quantitative phase tomography algorithm appears to be sensitive to inhomogeneities between the focal layers originating from incomplete flat fielding and a small line-to-line instability of the read-out gain. A simple mitigation is to correct pixel values for line-to-line and pixel-to-pixel variations by the average line and pixel value:

$$I'[l,p,c,z] = I[l,p,c,z] - \frac{1}{N_p}\sum_p I[l,p,c,z] - \frac{1}{N_l}\sum_l I[l,p,c,z]. \tag{3.17}$$

This can work because the DC offset is not important for the phase tomography algorithm.

3.3. Single scan multi-focal volumetric imaging of thick tissue layers

3.3.1. Axial deconvolution

Widefield microscopy has no optical sectioning capability, i.e. multi-focal volumetric images have optical cross-talk, adding blurred structures in out-of-focus layers to the in-focus image. This can be partly compensated by the use of deconvolution techniques, in case of sufficiently high SNR [45–47]. A drawback of applying existing 3D deconvolution approaches to multi-focal whole-slide images is the computational complexity, which scales unfavorably with lateral image size. This may be attributed to the lateral sharpening, that is a secondary goal of 3D deconvolution methods. The idea of this lateral sharpening is to partially undo the blurring by the microscope Point Spread Function (PSF). Here, a 'z-only' approach is followed, in which the goal of lateral sharpening is abandoned, and the focus is entirely on the goal of reducing the axial cross-talk. This reduces the computational complexity in such a way that the computational costs are scalable with lateral image size in a straightforward way. Deconvolution would then become suitable for real-time processing with the use of parallel processing.

The deconvolution algorithm is a variant of the Iterative Constrained Tikhonov-Miller (ICTM) deconvolution method [48], which is based on the minimization of the function:

$$\epsilon = \frac{1}{2}\left|p - Gd\right|^2 + \frac{1}{2}w\left|Cd\right|^2 \tag{3.18}$$

under the constraint $d \geq 0$. Here p represents the pre-processed image data vector, d the to-be-deconvolved image data vector, G the blurring matrix (PSF), w the regularization parameter, and C the regularization matrix. The constraint $d \geq 0$ is taken into account by a mapping $d = c(s)$ from the domain $-\infty < s < \infty$ to $0 \leq d < \infty$. Verveer and Jovin [48] use the mapping $c(s) = s$ for $s \geq 0$ and $c(s) = 0$ for $s < 0$. The conjugate gradient method is used in ref. [48] to minimize Eq. (3.18) with the aid of this mapping. An alternative mapping that could possibly work is $c(s) = \exp(s)$. The gradient vector of the error function w.r.t. d has components:

$$r_i = \frac{\partial \epsilon}{\partial d_i} = \sum_j A_{ij} d_j - b_i, \tag{3.19}$$

where i and j are indices labeling the vector and matrix components, and with the

vector $b = G^T p$ and the matrix $A = G^T G + wC^T C$. The Hessian is:

$$\frac{\partial^2 \epsilon}{\partial d_i \partial d_j} = A_{ij}. \tag{3.20}$$

For the unconstrained vector s we find a gradient:

$$u_i = \frac{\partial \epsilon}{\partial s_i} = c'(s_i)\, r_i \tag{3.21}$$

and a Hessian:

$$B_{ij} = \frac{\partial^2 \epsilon}{\partial s_i \partial s_j} = c'(s_i)\, c'\left(s_j\right) A_{ij} + c''(s_i)\, r_i \delta_{ij}, \tag{3.22}$$

where $c'(s)$ and $c''(s)$ are the first and second order derivative of the mapping function. For the mapping function of ref. [48] $c'(s) = 1$ for $s > 0$ and $c'(s) = 0$ for $s \leq 0$, and $c''(s) = 0$ (the delta-peak at $s = 0$ is ignored). This boils down to $u = Pr$ and $B = P^T AP$ with P a diagonal (projection) matrix with entries $P_{ii} = 1$ for $s_i > 0$ and $P_{ii} = 0$ for $s_i \leq 0$. An iterative update is made of the unconstrained deconvolved image vector s at iteration step k in the search direction v following:

$$s^{k+1} = s^k + \alpha v \tag{3.23}$$

with step size:

$$\alpha = -\frac{u^T v}{v^T B v}. \tag{3.24}$$

The search direction can be computed using e.g. the conjugate gradient method or the steepest descent method. For the sake of simplicity we use the gradient in the original d-domain, i.e. $v = r$. The constrained deconvolved image vector d at iteration $k+1$ is then found by $d^{k+1} = c\left(s^{k+1}\right)$. A fixed number of iterations N_i is used, because this removes the need for convergence testing during the optimization, which greatly contributes to the computational efficiency. For $N_i = 30$, the deconvolved images are converged to the final image with an error below the eight bit precision level. The ICTM method is applied to do only an axial deconvolution. This 'z-only' approach uses an effective blurring matrix:

$$G_{ij} = \alpha_j \left[1 - \beta + \beta \operatorname{sinc}\left(\pi \frac{z_i - z_j}{\Delta z}\right)^2\right], \tag{3.25}$$

of size $N_s \times N_s$ with N_s the number of layers, $\operatorname{sinc}(t) = \sin(t)/t$, and z is the axial position in object space. The constant term represents the axial response for a uniform object (delta-peak in Fourier space), the sinc term represents the axial response for a point object (uniform in Fourier space). The weighting coefficient β takes the relative magnitude of both contributions into account. In practice, a value close to but not equal to one gives satisfactory results. In the following $\beta = 0.95$ has been used. The parameter Δz is a measure for the depth of focus for which we use $\Delta z = 19.9\,\mu\text{m}$ The parameter α_j is used to normalize every row of the blurring matrix G such that $\sum_j G_{ij} = 1$. This ensures that G leaves the DC component unchanged, and it prevents the deconvolution to underestimate the value of d at the upper and lower scanned layers.

Regularization is needed in order to avoid division by zero if the blurring matrix is singular and/or if prior knowledge of the imaged object is available. For the axial deconvolution with the effective blurring matrix Eq. (3.25) this appears to be relatively unimportant. In the following $C = I$, the identity matrix, is used and $w = 10^{-6}$.

The axial deconvolution method is implemented on a GPU using MATLABs Parallel Processing Toolbox. Deconvolving an image segment of $512 \times 4096 \times 3 \times 8$ requires 17 ms on an Nvidia Tesla P100-PCIE-16GB GPU. The current calculation time implies a throughput of about 3 GPx/s, far exceeding the acquisition speed.

3.3.2. Scan setup and samples

The 10× objective lens is used to cover a large axial range for multi-focal volumetric imaging of thick slides. A sensorlet pitch $N_t = 18$ is used such that the layers have a pitch of $N_t\Delta_{\text{ax}} = 11.7\,\mu\text{m}$ and cover a total axial range of $8N_t\Delta_{\text{ax}} = 93.5\,\mu\text{m}$

Two samples are used to demonstrate the imaging of thick slides. The first slide contains a 60 µm thick stage 3 human rectum cancer section, stained with Hematoxylin and Eosin (H&E). The second slide contains a 60 µm thick human prostate section, stained with H&E and cleared using benzyl alcohol benzyl benzoate (BABB) [49].

3.3.3. Results

Figure 3.4 shows the result of a multi-focal scan of a normal human rectum tissue section. It provides a side-by-side comparison of the image after pre-processing and the image after deconvolution. In Figs. 3.4a and 3.4b, the center layer of the focal stack is shown, displaying a $800 \times 800\,\mu\text{m}^2$ area of 60 µm thick mucosa tissue. Transverse sectioned colonic crypts are shown (tube shaped mucosal structures) extended towards the bowel lumen. An edge towards the lumen is shown in the left half of the image, see Arrow 1 in Fig. 3.4. For two structures, a detailed view of all focal layers is provided in Figs. 3.4c to 3.4f.

The deconvolution algorithm clearly improves the image contrast by using the information of the neighboring layers. A higher contrast between the dark ring of cells along the crypts and the surrounding area is observed, see e.g. Arrow 2 in Figs. 3.4a and 3.4b. Also, the nuclei in the intermediate tissue show enhanced contrast, see e.g. Arrow 3 in Figs. 3.4a and 3.4b. The deconvolution reveals multiple cellular layers of the cellular inner structure of the crypt that are hard to distinguish in the pre-processed data, see e.g. Arrow 4 in Figs. 3.4c and 3.4d and Arrow 5 in Figs. 3.4e and 3.4f. Some ringing artifacts are observed in in the out-of-focus layers, in particular for $z = 47\,\mu\text{m}$, see e.g. Arrow 6 in Figs. 3.4c and 3.4d and Arrow 7 in Figs. 3.4e and 3.4f, but do not seem to harm the overall image quality too much.

The result of a multi-focal scan of a cleared prostate section is provided in Fig. 3.5. Again, a side by side comparison is provided between the pre-processed data and the deconvolved data. Figs. 3.5a and 3.5b depict an $800 \times 800\,\mu\text{m}^2$ area of 60 µm thick prostate tissue, showing stroma separated from prostate glandules (containing two types of cells in normal prostate). In Figs. 3.5c and 3.5d a detailed view is provided of an almost fully cleared tissue, where mainly the nuclei have remained visible. Figures 3.5e and 3.5f provide a detailed view of the differences in morphology.

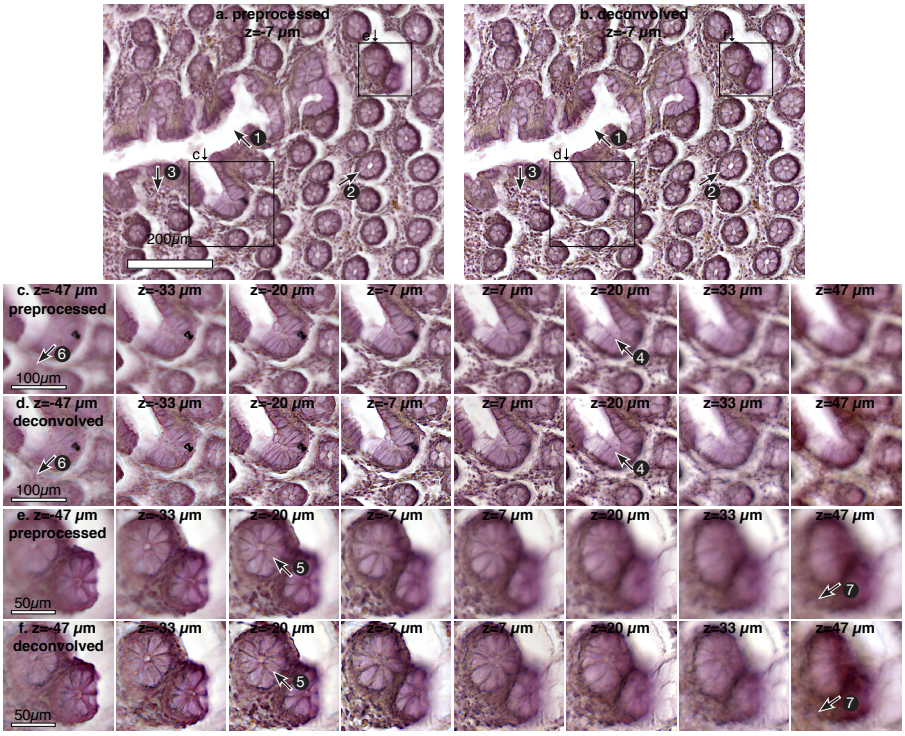


Figure 3.4: A multi-layer scan of a 60 µm thick, H&E stained normal human rectal mucosa section. (a) Center layer after pre-processing. (b) Center layer after deconvolution. Visualization 1* provides the whole focal volume. (c, e) All focal layers of a detail of the pre-processed image. (d, f) All focal layers of a detail of the deconvolved image.
* doi.org/10.6084/m9.figshare.12097872.v1

The deconvolved images show an improved contrast, especially on the smaller length scales. For example, the nuclei show a good separation, see e.g. Arrow 1 in Figs. 3.5c and 3.5d. Also there is a clear separation of basal and luminal cells throughout the focal volume, see e.g. Arrow 2 in Figs. 3.5e and 3.5f. Bright circular shapes with a smooth gray outline are visible in Fig. 3.5d for $z \leq -33\,\mu\text{m}$, see for example Arrow 3 in Fig. 3.5d These structures are out-of-focus nuclei of which only the center is suppressed by the axial deconvolution. This is interpreted as an artefact of the 'z-only approach' to deconvolution.

3.4. Quantitative Phase Tomography of unstained tissue layers

3.4.1. Quantitative Phase Tomography

We have implemented 3D phase retrieval using the Quantitative Phase Tomography (QPT) method that was recently introduced by Descloux *et al.* [41]. This method enables the retrieval of a local phase in 3D from a stack of bright-field images by a simple filtering operation. Their optical analysis is based on the Born approximation, in which the 3D intensity of a weak scattering object can be writ-

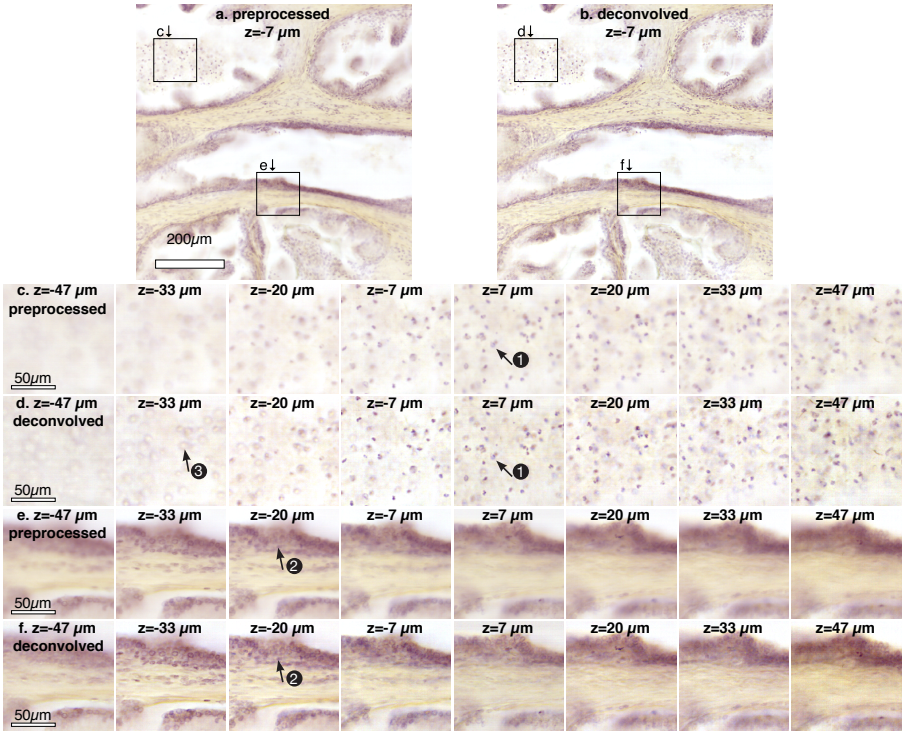


Figure 3.5: A multi-layer scan of a 60 µm thick, cleared, and H&E stained human prostate section. (a) Center layer after pre-processing. (b) Center layer after deconvolution. Visualization 2* provides a side-by-side comparison of the whole focal volume. (c, e) All focal layers of a detail of the pre-processed image. (d, f) All focal layers of a detail of the deconvolved image.
* doi.org/10.6084/m9.figshare.12097872.v1

ten in 3D Fourier space as a function of the 3D spatial frequency vector $\vec{f}$ as:

$$\hat{I}(\vec{f}) \approx I_0\delta(\vec{f}) + \hat{\Gamma}(\vec{f}) + \hat{\Gamma}^*(\vec{f}), \tag{3.26}$$

The first term represents a DC offset with amplitude I_0, giving rise to the delta-function $\delta(\vec{f})$ in reciprocal space. The second and third term are the complex valued so-called cross-spectral density $\hat{\Gamma}(\vec{f})$ and its complex conjugate.

After a 3D Fourier transform from spatial frequency to real space, the cross-spectral density is related to the local phase by:

$$\phi(\vec{r}) = \tan^{-1}\left(\frac{\alpha \mathrm{Im}(\Gamma(\vec{r}))}{I_0 + \alpha \mathrm{Re}(\Gamma(\vec{r}))}\right), \tag{3.27}$$

where α is a calibration factor and $\vec{r} = (x, y, z)$ is the spatial position vector.

Descloux *et al.* provide a description of the spatial frequency support of the cross-spectral density, which is given by the difference between any possible spatial frequency vector in the cone of directions of incidence on the illumination side and in the cone of directions of scattering on the detection side:

$$\vec{f} = \frac{n}{\lambda}\left(\begin{bmatrix}\sin\theta_i \\ \cos\theta_i\end{bmatrix} - \begin{bmatrix}\sin\theta_d \\ \cos\theta_d\end{bmatrix}\right), \text{with} \, |\theta_i| < \arcsin\frac{\mathrm{NA}_i}{n}, \text{and} \, |\theta_d| < \arcsin\frac{\mathrm{NA}_d}{n}, \tag{3.28}$$

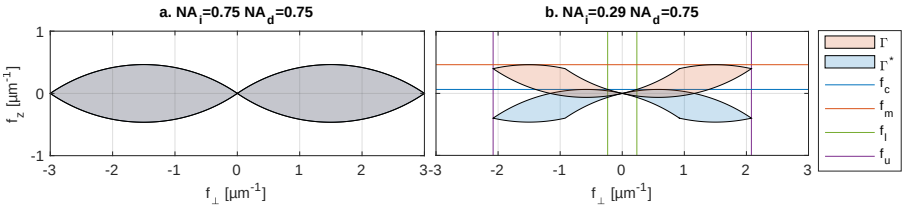


Figure 3.6: The frequency support of the cross-spectral density $\hat{\Gamma}$ and its complex conjugate $\hat{\Gamma}^*$ as a function of the lateral spatial frequency $f_\perp = \sqrt{{f_x}^2 + {f_y}^2}$ and the axial spatial frequency f_z for a wavelength $\lambda = 500\,\mathrm{nm}$. (a) In the incoherent limit $\mathrm{NA}_i = \mathrm{NA}_d$, the support of $\hat{\Gamma}$ fully overlaps with the support of $\hat{\Gamma}^*$. (b) Under partial coherence conditions $\mathrm{NA}_i < \mathrm{NA}_d$, $\hat{\Gamma}$ can be retrieved in the region $f_z > f_c$ and $f_l < |f_\perp| < f_u$ from the intensity data.

where n is the average object refractive index, NA_i is the illumination numerical aperture, NA_d is the detection numerical aperture, and λ the wavelength. Only the spatial frequency component along the radial direction ($f_\perp$) and along the optical axis (f_z) are given, as the optical system is rotationally symmetric around the optical (z) axis. The frequency support of the complex conjugate $\hat{\Gamma}^*$ has the same shape, but is mirrored in the $f_\perp$ axis (substitution $f_z \rightarrow -f_z$). Figure 3.6a shows the frequency support in the incoherent limit, i.e. the detection NA_d is equal to the illumination NA_i, where $\hat{\Gamma}$ fully overlaps with its complex conjugate. In contrast, in a partially coherent system, i.e. NA_i<NA_d, $\hat{\Gamma}$ can be largely separated from $\hat{\Gamma}^*$, see Fig. 3.6b. The frequency support of $\hat{\Gamma}$ has an upper bound in f_z given by:

$$f_m = \frac{n}{\lambda}\left(1 - \sqrt{1 - \frac{\mathrm{NA}_d^2}{n^2}}\right). \tag{3.29}$$

The upper bound in f_z of the complex conjugate Γ^* is:

$$f_c = \frac{n}{\lambda}\left(1 - \sqrt{1 - \frac{\mathrm{NA}_i^2}{n^2}}\right). \tag{3.30}$$

Both f_m and f_c are indicated in Fig. 3.6b. Here it can be seen that a substantial part of $\hat{\Gamma}$ can be retrieved from a stack of intensity data by using a single side-band high-pass filter K:

$$\hat{\Gamma}_+(\vec{f}) = \hat{I}(\vec{f})\hat{K}(\vec{f}),\ \hat{K}(\vec{f}) = \begin{cases} 1, & f_z > f_c \\ 0, & \text{otherwise} \end{cases}. \tag{3.31}$$

It can be seen from Eq. (3.28) that, provided that the NA_i substantially smaller than NA_d, the support of Γ is bound by the arc $(f_\perp \lambda/n + \sin\theta_i)^2 + (f_z\lambda/n - \cos\theta_i)^2 = 1$. The intersection of this curve with $f_z = f_c$ provides the smallest lateral spatial frequency f_l for which $\hat{\Gamma}_+$ is nonzero. This smallest lateral spatial frequency f_l is given by:

$$f_l = \frac{n}{\lambda}\left[2\sqrt{\sqrt{1 - \frac{\mathrm{NA}_i^2}{n^2}} - \left(1 - \frac{\mathrm{NA}_i^2}{n^2}\right)} - \frac{\mathrm{NA}_i}{n}\right] \tag{3.32}$$

The highest lateral spatial frequency for which $\hat{\Gamma}_+$ is nonzero is given by the regular partially coherent cutoff frequency:

$$f_u = \frac{\mathrm{NA}_d}{\lambda} + \frac{\mathrm{NA}_i}{\lambda}. \tag{3.33}$$

In Fig. 3.6b both f_l and f_u are indicated.

It may be concluded that, although we do not apply explicit filtering in the lateral direction, the Optical Transfer Function (OTF) of this new phase imaging modality will be a band pass filter in the lateral plane. The absence of low spatial frequencies is not uncommon to phase reconstructions obtained from through-focus image stacks. Methods based on solving the Transport of Intensity Equation (TIE) [35–38] need to invert the Laplacian in the lateral coordinates. The Laplacian has a transfer function that depends quadratically on the (lateral) spatial frequency components, i.e. it has (near) zero transfer at low spatial frequencies. The band-pass character of the QPT modality under consideration will result in phase reconstructions in which there is considerable edge ringing. Two uniform regions separated by a sharp phase step are recognizable as flat regions with the same phase value, the border between the two regions separated by a single oscillation. A small point-like phase object is imaged as a λ/NA_d sized spot with considerable circular fringes. In fact the integral over the entire Point Spread Function (PSF) must be zero, as the transfer function at zero spatial frequency is zero.

In the following we present an analysis of different trade-offs between and interdependencies of the optical system parameters. There are six system parameters for quantitative phase imaging: the NA of the detection NA_d, the NA of the illumination NA_i, the average sample refractive index n, the wavelength λ, the axial sampling distance Δz, and the number of layers scanned N_s. These parameters, however, are not independent. First, Nyquist sampling in the axial direction requires an axial sampling:

$$\Delta z = \frac{1}{2f_m} = \frac{\lambda}{2n}\left(1 - \sqrt{1 - \frac{\mathrm{NA}_d^2}{n^2}}\right)^{-1}, \tag{3.34}$$

where we used the relation between f_m and NA_d given in Eq. (3.29). Second, it appears that there is an optimal choice for the illumination NA_i given the number of scanned layers N_s. Equation (3.30) implies that the illumination NA_i can be derived from the lower axial cutoff frequency f_c. Generally, a lower f_c is favorable, but f_c is limited by the sampling density in the Fourier domain. For that reason f_c is selected to be half of the smallest resolvable frequency:

$$f_c = \frac{1}{2N_s\Delta z}. \tag{3.35}$$

The corresponding NA_i can be found by solving Eq. (3.30) and substituting Eq. (3.35) to obtain:

$$\mathrm{NA}_i = n\sqrt{1 - \left(1 - \frac{1 - \sqrt{1 - \mathrm{NA}_d^2/n^2}}{N_s}\right)^2}. \tag{3.36}$$

Substitution in Eq. (3.30) leads to:

$$f_c = \frac{f_m}{N_s}, \tag{3.37}$$

i.e. the ratio to the upper and lower axial cutoff frequencies must be equal to the number of scanned layers. It is noted that the system has no direct dependence on n, but only on the ratio's NA_d/n and NA_i/n. The current analysis shows that instead of six, there are only three degrees of freedom in the design of the optical system. A convenient set of three independent system parameters are formed by the diffraction length scale λ/NA_d, the scaled objective lens NA_d/n, and the number of scanned layers N_s.

In Fig. 3.7 different relevant optical system parameters are plotted as a function of N_s for four different NA_d/n values corresponding to objective lenses with NA_d = 0.2, 0.45, 0.75, 1.2 and an average sample refractive index n = 1.33. Figure 3.7a shows the axial spatial frequency support of $\hat{\Gamma}_+$, i.e. the range of spatial frequencies between f_c and f_m. The maximum axial spatial frequency f_m is independent of the number of layers scanned but shows a strong dependence on NA_d/n. This implies that the axial resolution depends on NA comparable to conventional brightfield or fluorescence microscopes. The lower bound of the spatial frequency support f_c appears to be inversely proportional to the number of scanned layers N_s according to Eq. (3.37). Imaging objects with large axial size, i.e. small axial spatial frequencies, is thus realized primarily by scanning more layers.

The lateral frequency support of $\hat{\Gamma}_+$, i.e., the range of frequencies between the lower and upper cutoff spatial frequencies f_l and f_u is shown in Fig. 3.7b. The lateral frequency support depends only slightly on NA_d/n. For increasing values of N_s, the lateral upper bound f_u decreases to an asymptotic value λ/NA_d. The lateral lower bound f_l decreases typically as $1/\sqrt{N_s}$, except for the practically not so relevant case N_s = 2 or 3 in combination with a high value of NA_d/n. Imaging objects with large lateral size, i.e. small lateral spatial frequencies, can thus be improved by scanning more layers, just as for axial case, although the rate of improvement scales less favourably for the lateral case.

Figure 3.7c shows the partial coherence factor $\mathrm{NA}_i/\mathrm{NA}_d$, which turns out to depend hardly on NA_d/n. The partial coherence factor decreases with N_s typically as $1/\sqrt{N_s}$.

3.4.2. Numerical implementation

The QPT method takes the following steps [41]: 1. Padding the intensity data $I(x,y,z)$ with a mirrored copy $I(x,y,-z)$ to minimize Fourier streaking occurring because of the boundary discontinuity in the axial direction. 2. Take a 3D Fourier transform of the data. 3. Apply the axial spatial frequency filter. 4. Optionally, noise can be reduced by suppressing all spatial frequencies where no signal is expected based on the known spatial frequency support given in Eq. (3.28). 5. Take the inverse 3D Fourier transform. 6. Remove the padded data. 7. Calculate the phase using Eq. (3.27).

In our application, scalability to large image datasets is of utmost importance. We therefore do not implement the optional 4th step. This makes the implementation a z-only problem, just as for the deconvolution approach we propose, easily

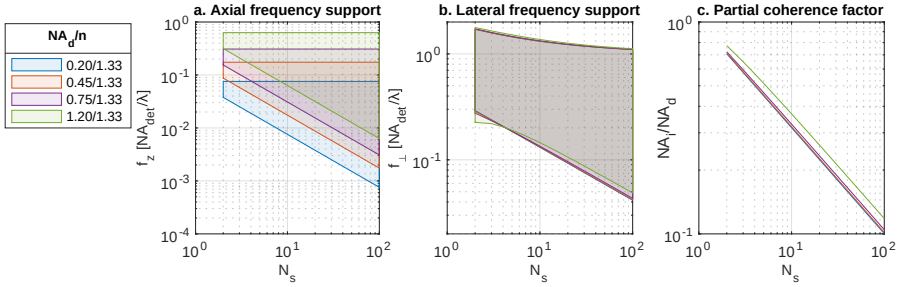


Figure 3.7: The behavior of a QPT optical system expressed in normalized coordinates, as a function of the dimensionless numbers N_s and NA_d/n. (a) The axial frequency support of $\hat{\Gamma}_+$. (b) The lateral frequency support $\hat{\Gamma}_+$. (c) The partial coherence factor.

compatible with parallel processing solutions. It turns out that is numerically convenient to implement steps 1-3, 5, and 6 using a single matrix transform:

$$\Gamma_+ = PI(\vec{r}) = \left[CF_z^{-1}KF_zM\right]I(\vec{r}), \tag{3.38}$$

where P is a $N_s \times N_s$ matrix with N_s the number of layers in the intensity stack $I(\vec{r})$. P is composed of four matrices. First, M is a $2N_s \times N_s$ matrix that pads the data using $M_{ij} = \delta_{ij} + \delta_{2N_s-i+1,j}$ where δ is the Kronecker delta. The second matrix F has size $2N_s \times 2N_s$ and implements the Fourier transform along the z-axis. Matrix K is a $2N_s \times 2N_s$ diagonal matrix implementing the spatial frequency filter. The data is inverse Fourier transformed using the $2N_s \times 2N_s$ matrix F^{-1}. Finally, the mirrored data is removed by the $N_s \times 2N_s$ matrix $C_{ij} = \delta_{ij}$ for $i, j \leq N_s$ and zero otherwise. The matrix P has to be calculated only once and can then be applied to every coordinate (x,y) independently. This method is particularly suitable for small numbers of layers, where the need for extra memory allocation that is associated with data mirroring outweighs the high efficiency of the Fast Fourier Transform compared to a matrix multiplication. This method is implemented for execution on a Graphics Processing Unit (GPU) using the Matlab Parallel Processing Toolbox. Calculating the phase of a $2048 \times 2048 \times 8$ image requires about 50 ms using an Nvidia Tesla P100-PCIE-16GB. This implies a throughput of 670 MPx.

3.4.3. Scan setup and samples

The QPT method is implemented based on scans with the 20× objective lens, the green color channel, and the use of $N_s = 8$ scan lines simultaneously. Given that $\mathrm{NA}_d = 0.75$, and assuming a sample refractive index of $n = 1.33$ and a wavelength of $\lambda = 500\,\mathrm{nm}$, Eq. (3.34) results in a target Δz of 1.05 µm. To guarantee Nyquist sampling a sensorlet interval $N_t = 6$ is chosen, leading to $\Delta z = 0.93\,\mu\mathrm{m}$. This gives a total covered axial range $N_s\Delta z = 7.4\,\mu\mathrm{m}$. Now Eq. (3.35) provides $f_c = 1/14.9\,\mu\mathrm{m}$, and solving Eq. (3.30) gives $\mathrm{NA}_i = 0.29$. As a result, the system will have a lateral resolution of $1/f_u = 0.48\,\mu\mathrm{m}$, as follows from Eq. (3.33), it images structures with a lateral size up to $1/f_l = 4.3\,\mu\mathrm{m}$, according to Eq. (3.32), and Eq. (3.29) provides a maximum axial spatial frequency $f_m = 1/2.2\,\mu\mathrm{m}$.

Three samples are used to demonstrate the QPT method. The first slide contains a 5 µm thick human prostate Tissue Micro Array (TMA) section labeled with

Kreatech ERBB2 (17q12) / SE 17 FISH probe (product number KBI-10701) for the detection of amplification of the ERBB2 (also known as HER-2/*neu*) gene via Fluorescence *in situ* Hybridization (FISH) [50]. The second slide is a 4 µm thick stage 3 human rectum cancer section with immunofluorescence labeling. Three proteins are labeled: Desmin (IgG1 M antibody labeled with Alexa Fluor 488), which is highly expressed in muscle cells, CD31 (IgG R antibody labeled with Alexa Fluor 546), which is a is a marker for blood vessels, and D2-40 (labelled with Alexa Fluor 594), which is used as a marker of lymphatic endothelium. Additionally, the slide is stained with DAPI, labelling the nuclei. The third slide is an 4 µm thick human prostate section. The slide is deparaffinized and embedded in xylene but not further processed for staining.

3.4.4. Results

Figure 3.8 shows the result for the ERBB2 slide. Figure 3.8a shows all eight layers of the pre-processed image data. A detail is selected in which a single cell is visible. The local phase computed from this data is shown in Fig. 3.8b. A diverging colormap is used to display the phase values [52], with blue corresponding to negative values, black to zero and green to positive values. The total estimated optical thickness of the sample is obtained by summing the local phase over all layers. The result is shown over three length scales, zooming in with a factor of 64: Fig. 3.8c has the full width of a single scan lane, Fig. 3.8d displays an intermediate length scale, and Fig. 3.8e corresponds to the detail shown in Figs. 3.8a and 3.8b. Figures 3.8f to 3.8h show a color coded maximum intensity projection of the same area. The pixel values in this image have an intensity corresponding to the maximum phase along the axial direction and a color corresponding to the depth at which the maximum was found. The used colormap, shown at the left side of Fig. 3.8f, has a uniform luminescence and an equidistant color spacing. This minimizes the visual cross-talk between depth and intensity. As a reference, Figures 3.8i to 3.8k provide a maximum intensity projection of a multi-focal fluorescence image of the same section, for the same areas. This fluorescence image was obtained in previous research [27] using a multi-focal multi-line confocal scanner. A rigid transform is used to register the fluorescence reference image to the phase image, where the optimal transform is found by minimizing the root mean square distance between a series of manually selected landmarks.

In the pre-processed images shown in Fig. 3.8a the outline of the cells are visible at low contrast, as well as a few spots that change from bright to dark through focus, see Arrows 1 and 2 in Fig. 3.8a. The corresponding local phase in Fig. 3.8b shows the contours of the cells and cell organelles with a high contrast. The QPT algorithm is able to reveal the axial position ($z \approx 1.7\,\mu$m) of the sources of the bright and dark spots, see Arrow 3 in Fig. 3.8c. The lateral band-pass behavior of the QPT algorithm causes ringing, which is particularly visible as a blue (negative) local phase surrounding the cell, see for example Arrow 4 in Fig. 3.4b. Also in the total optical thickness, displayed in Figs. 3.8c to 3.8e, the cell contours and tissue structure is clearly revealed. The band-pass characteristic is again apparent, for example indicated by Arrow 5 in Fig. 3.8d and Arrow 6 in Fig. 3.8e. The color coded maximum intensity projection appears particular suitable for imaging at larger length scales. For example, the wall of a blood vessel with a size of $\sim 100\,\mu$m is clearly vis-

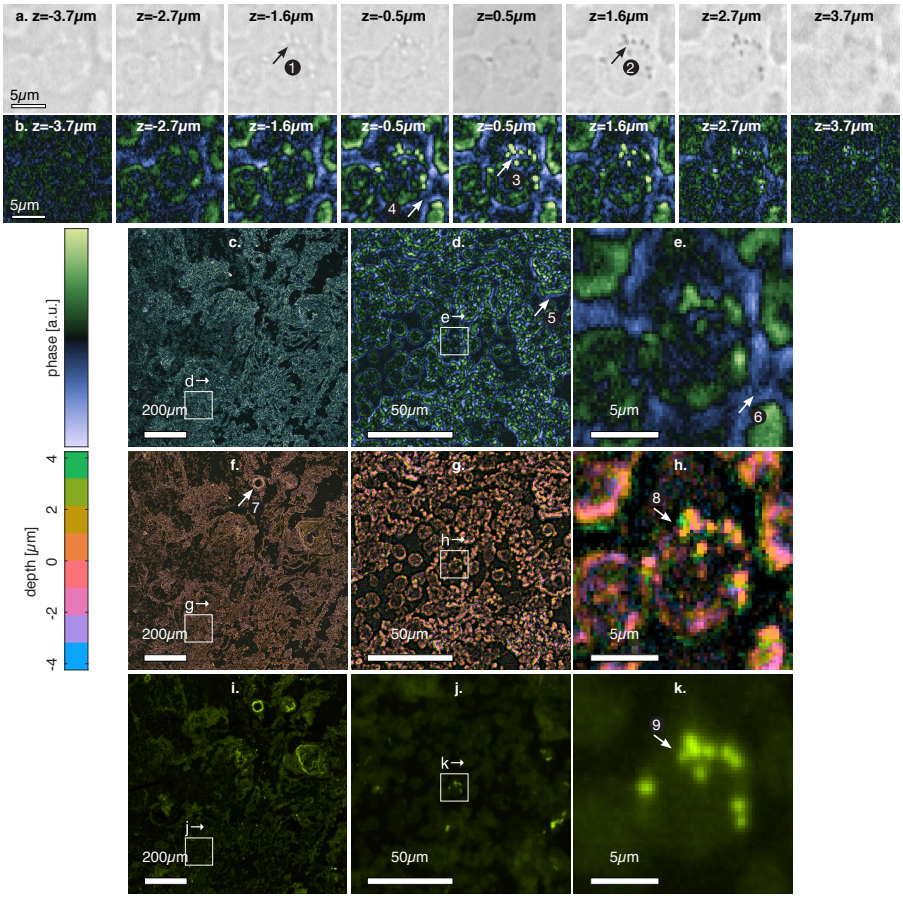


Figure 3.8: A multi-layer phase contrast image of a 5 µm thick human prostate TMA section prepared for ERBB2 detection using FISH . The layers have an axial spacing of 0.93 µm spanning 7.4 µm in total. (a) A detail of the raw image data after pre-processing. (b) The retrieved phase of the corresponding area. (c-e) The total optical thickness, shown over three length scales. Visualization 3[a] shows this in more detail. (f-h) A color coded maximum intensity projection of the same area, see Visualization 4[b]. (i-k) A confocal fluorescence image of the same section [27] showing the same area, see Visualization 5[c].

[a] doi.org/10.6084/m9.figshare.12097881.v1
[b] doi.org/10.6084/m9.figshare.12097884.v1
[c] doi.org/10.6084/m9.figshare.12097887.v1

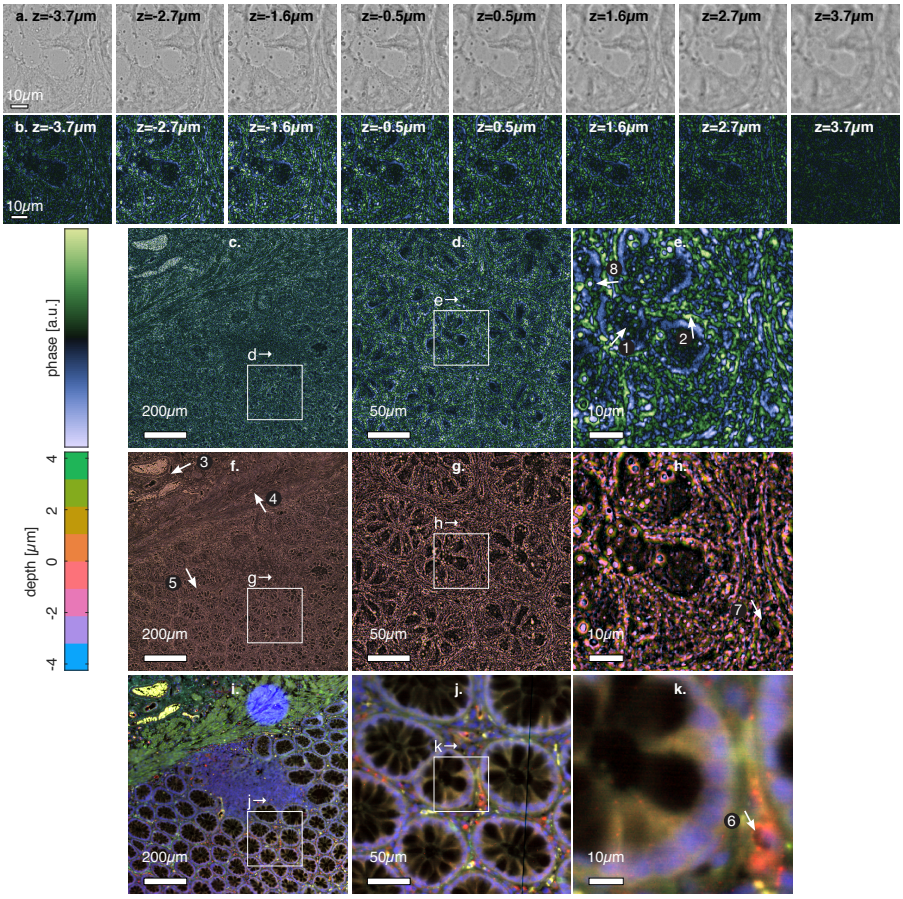


Figure 3.9: A multi-layer phase contrast image of a 4 µm thick human rectum section. The layers have an axial spacing of 0.93 µm spanning 7.4 µm in total. (a) A detail of the raw image data after preprocessing. (b) The retrieved phase of the corresponding area. (c-e) The total optical thickness, shown over three length scales. Visualization 6[a] shows this in more detail. (f-h) A color coded maximum intensity projection of the same area, see Visualization 7[b]. (i-k) A widefield multi-color fluorescence image of an directly adjacent section [51] showing the same area, see Visualization 8[c].

[a] doi.org/10.6084/m9.figshare.12097890.v1
[b] doi.org/10.6084/m9.figshare.12097893.v1
[c] doi.org/10.6084/m9.figshare.12097896.v1

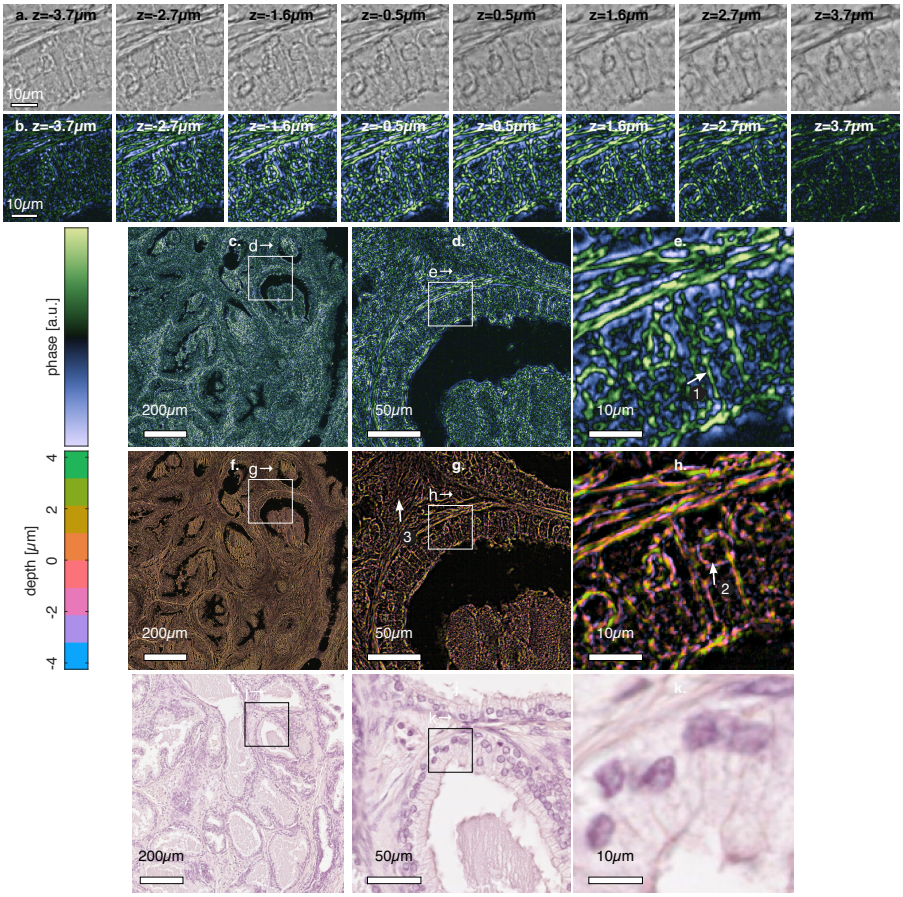


Figure 3.10: A multi-layer phase contrast image of an unstained, 4 µm thick human prostate section. The layers have an axial spacing of 0.93 µm spanning 7.4 µm in total. (a) A detail of the raw image data after pre-processing. (b) The retrieved phase of the corresponding area. (c-e) The total optical thickness, shown over three length scales. Visualization 9[a] shows this in more detail. (f-h) A color coded maximum intensity projection of the same area, see Visualization 10[b]. (i-k) An H&E stained section of the same tissue block showing the same area, see Visualization 11[c].

[a] doi.org/10.6084/m9.figshare.12097899.v1
[b] doi.org/10.6084/m9.figshare.12097902.v1
[c] doi.org/10.6084/m9.figshare.12097905.v1

ible, see Arrow 7 in Fig. 3.8f. Comparing the results with the fluorescence images, a surprisingly clear correspondence is found between the cell organelles observed in the phase images and the FISH labeled sites, compare Arrow 8 in Fig. 3.8h with Arrow 9 in Fig. 3.8k.

Figure 3.9 shows the result for the human rectum section following the same structure as Fig. 3.8. The sample is shown over three length scales, zooming in with a factor of 16, ranging from the full width of a single scan lane to the same detail shown in Figs. 3.9a and 3.9b. Figures 3.9i to 3.9k provide a multi-color wide-field fluorescence image of a directly adjacent, identically prepared slide that we reported on in earlier research [51]. This image was registered to the phase image, just as for the previous case.

The insets show the cross section of normal crypts in the human rectum. The crypt lumen indicated by Arrow 1 Fig. 3.9e and cell walls indicated by Arrow 2 Fig. 3.9e are clearly visible. The phase images are able to reveal relevant structure in the tissue on the larger length scales as well, in particular in the color coded maximum intensity projection. For example, the top half of the largest zoom level shows the submucosa with two veins, see Arrow 3 in Fig. 3.9f, the muscularis mucosae, see Arrow 4 in Fig. 3.9f, while the bottom half shows the normal mucosa with the crypts, indicated by Arrow 5 in Fig. 3.9f. A clear correspondence is found with the fluorescence images. For example, the red labeled micro vessel indicated by Arrow 6 in Fig. 3.9h can also be observed in the phase contrast images, see Arrow 7 in Fig. 3.9h. At the smallest zoom level small spots are visible that are not present in the fluorescence control images. Interestingly, some of them have a negative phase value, see e.g. Arrow 8 in Fig. 3.9e, suggesting that this might be regions with a lower refractive index than the surrounding tissue structures, such as water droplets or air bubbles. Around the spots, ringing is visible in correspondence with the expected bandpass behavior.

Figure 3.10 shows the result for the human prostate section following the same structure as Fig. 3.8 and Fig. 3.9. The result is shown over three length scales, zooming in over a factor of 25, from the full width of a single scan lane down to a columnar epithelium layer around a lumen. As a reference an H&E stained section from the same tissue block is shown, registered to the phase image. It is noted, however, that this section was not directly adjacent to the unstained section used for phase imaging, and that therefore the overall structure is not corresponding closely to the structure of the phase images of the unstained slide.

The cell borders of the epithelium are clearly visible in the phase images, see Arrow 1 in 3.10e. Also apparent are the outline of the nuclei e.g. as indicated by Arrow 2 in 3.10h and the structure of the stroma e.g. as indicated by Arrow 3 in 3.10g. The overall features (highlighting edges and near-point like objects) are the same as for the other two cases.

3.5. Conclusion

In conclusion, we present a whole slide imaging (WSI) system based on a multi-line CMOS sensor devised by Philips [24–26]. A tilted image plane makes it possible to acquire image data from multiple focal slices simultaneously. The architecture with a single image sensor provides inherent registration of different color channels and focal layers. The ‘push-broom’ scanning approach results in a field of

view that is in principle unlimited, and a reduced need for stitching. The system achieves image acquisition with a very high throughput of ~400 MPx/s, and can be used for 2D full-colour imaging of absorption stained slides with continuous autofocus, but also for multi-focus imaging.

The scanner platform enables novel contrast modalities based on computational imaging approaches. Multi-focal volumetric imaging of thick pathology samples is demonstrated, where 8 layers are acquired in a single scan. Samples of 60 µm thick are imaged with a resolution of 0.5 µm. Deconvolution is used to improve contrast, which is inherently low for such thick specimens. A simplified form of the ICTM deconvolution method is proposed, targeting suppression of out-of-focus light only, and ignoring lateral resolution improvement, enabling very high processing speeds of about 3 GPx/s, far exceeding the acquisition speed. This indicates that this approach is suitable for on-line, and potentially on-chip, processing.

The sectioning of the focal layers could possibly be improved upon by combining the proposed 'z-only' deconvolution with a multi-scale image approach [53]. Decomposition of the image data in a wavelet representation would enable a different setting of the depth of focus parameter for every lateral length scale in the wavelet domain [54]. In this way, the deconvolution would become effectively dependent on the lateral spatial frequency content, while at the same time avoiding a (much) larger blurring matrix.

Next, phase imaging based on the recently introduced QPT method [41] is demonstrated. A simplified form of the algorithm, based on 'z-only' processing is proposed for this modality as well, giving data processing speeds of ~0.67 GPx/s, exceeding the acquisition speed. A system design study of the phase imaging modality is developed in this paper, leading to a description in which the axial and lateral spatial frequency support, as well as the partial coherence factor, are entirely given by the diffraction length scale λ/NA_d, the scaled imaging NA_d/n and the number of scanned layers N_s. The image formation theory points to an in-plane transfer function that has the character of a band-pass spatial frequency filter. The lower and upper cutoff spatial frequencies for the current setup (with objective lens $\mathrm{NA}_d = 0.75$ and condenser $\mathrm{NA}_i = 0.26$) are at $1/f_l = 4.3$ µm and $1/f_u = 0.48$ µm. This band-pass transfer results in phase images where the near-point like objects, with a size of ~ 1 µm, are highlighted, and where phase step edges are recognizable by a single oscillation in the direction orthogonal to the edges. Comparison of QPT images of a slide prepared for FISH to the fluorescence reference image reveals that QPT imaging is able to image the sites labeled for FISH imaging. QPT imaging of an unstained tissue slide and comparison to immunofluorescence imaging shows that phase imaging can provide additional structural tissue information.

A next step for the QPT method could be to include a form of in-plane image processing, targeting to overcome the primary limitation of the QPT method of having a zero transfer function at low spatial frequencies, a limitation shared with TIE-based methods of solving the phase from through-focus image stacks. This next step should be accompanied with a careful balancing of signal reconstruction and noise amplification at low spatial frequencies, following the lines of e.g. [37, 38] for TIE-based methods. At the same time, sacrifices to computational speed should not be too large, as speed is a need for real-time processing of ~cm^2 tissue

areas scanned in ~1 min.

Funding
Nederlandse Organisatie voor Wetenschappelijk Onderzoek (Netherlands Organization forScientific Research - Applied and Engineering Sciences, NWO-TTW) (Project 13892).

Acknowledgments
We wish to thank Bas Hulsken, Rik Kneepkens, Jelte Peter Vink, Mathijs Rem, Gert-Jan van den Braak and Aslihan Arslan Carisey of Philips Digital & Computational Pathology, Hein Sleddens of the Erasmus University Medical Center, and Hans Morreau of the Leiden University Medical Center for valuable research input.

Disclosures
Philips Digital & Computational Patholog has made their prototype scanner infrastructure available to us for this research.

References

[1] L. van der Graaff et al. "Computational imaging modalities for multi-focal whole-slide imaging systems". In: *Applied Optics* 59.20 (2020), pp. 5967–5982. DOI: 10.1364/AO.394290 (cit. on p. 53).

[2] L. Pantanowitz et al. "Validating Whole Slide Imaging for Diagnostic Purposes in Pathology: Guideline from the College of American Pathologists Pathology and Laboratory Quality Center". In: *Archives of Pathology & Laboratory Medicine* 137.12 (2013), pp. 1710–1722. DOI: 10.5858/arpa.2013-0093-CP (cit. on p. 54).

[3] R. Randell, R. A. Ruddle, and D. Treanor. "Barriers and facilitators to the introduction of digital pathology for diagnostic work." In: *Studies in health technology and informatics* 216 (2015), pp. 443–447. DOI: 10.3233/978-1-61499-564-7-443 (cit. on p. 54).

[4] D. R. J. Snead et al. "Validation of digital pathology imaging for primary histopathological diagnosis". In: *Histopathology* 68.7 (2016), pp. 1063–1072. DOI: 10.1111/his.12879 (cit. on p. 54).

[5] E. Abels and L. Pantanowitz. "Current state of the regulatory trajectory for whole slide imaging devices in the USA". In: *Journal of Pathology Informatics* 8.1 (2017), p. 23. DOI: 10.4103/jpi.jpi_11_17 (cit. on p. 54).

[6] P. W. Hamilton, Y. Wang, and S. J. McCullough. "Virtual microscopy and digital pathology in training and education". In: *APMIS* 120.4 (2012), pp. 305–315. DOI: 10.1111/j.1600-0463.2011.02869.x (cit. on p. 54).

[7] K. Foster. "Medical education in the digital age: Digital whole slide imaging as an e-learning tool". In: *Journal of Pathology Informatics* 1.1 (2010), p. 14. DOI: 10.4103/2153-3539.68331 (cit. on p. 54).

[8] L. Pantanowitz and A. V. Parwani. "Education". In: *Digital Pathology*. Springer International Publishing, 2016, pp. 71–78. DOI: 10.1007/978-3-319-20379-9_7 (cit. on p. 54).

[9] E. A. Krupinski, A. K. Bhattacharyya, and R. S. Weinstein. "Telepathology and Digital Pathology Research". In: *Digital Pathology*. Springer International Publishing, 2016, pp. 41–54. DOI: 10.1007/978-3-319-20379-9_5 (cit. on p. 54).

[10] P. W. Hamilton et al. "Digital pathology and image analysis in tissue biomarker research". In: *Methods* 70.1 (2014), pp. 59–73. DOI: 10.1016/j.ymeth.2014.06.015 (cit. on p. 54).

[11] B. Hulsken, D. Vossen, and S. Stallinga. “High NA diffractive array illuminators and application in a multi-spot scanning microscope”. In: *Journal of the European Optical Society: Rapid Publications* 7 (2012), p. 12026. DOI: 10.2971/jeos.2012.12026 (cit. on p. 54).

[12] J. Wu et al. “Wide field-of-view microscope based on holographic focus grid illumination”. In: *Optics Letters* 35.13 (2010), pp. 2188–2190. DOI: 10.1364/OL.35.002188 (cit. on p. 54).

[13] S. Pang et al. “Wide and scalable field-of-view Talbot-grid-based fluorescence microscopy”. In: *Optics Letters* 37.23 (2012), pp. 5018–5020. DOI: 10.1364/OL.37.005018 (cit. on p. 54).

[14] G. Zheng, R. Horstmeyer, and C. Yang. “Wide-field, high-resolution Fourier ptychographic microscopy”. In: *Nature Photonics* 7.9 (2013), pp. 739–745. DOI: 10.1038/nphoton.2013.187 (cit. on p. 54).

[15] K. Guo, S. Dong, and G. Zheng. “Fourier Ptychography for Brightfield, Phase, Darkfield, Reflective, Multi-Slice, and Fluorescence Imaging”. In: *IEEE Journal of Selected Topics in Quantum Electronics* 22.4 (2016), pp. 77–88. DOI: 10.1109/JSTQE.2015.2504514 (cit. on p. 54).

[16] L. Tian et al. “Computational illumination for high-speed in vitro Fourier ptychographic microscopy”. In: *Optica* 2.10 (2015), pp. 904–911. DOI: 10.1364/OPTICA.2.000904 (cit. on p. 54).

[17] F. Ghaznavi et al. “Digital Imaging in Pathology: Whole-Slide Imaging and Beyond”. In: *Annual Review of Pathology: Mechanisms of Disease* 8.1 (2013), pp. 331–359. DOI: 10.1146/annurev-pathol-011811-120902 (cit. on pp. 54, 55).

[18] J. R. Gilbertson et al. “Primary histologic diagnosis using automated whole slide imaging: a validation study”. In: *BMC Clinical Pathology* 6.1 (2006), p. 4. DOI: 10.1186/1472-6890-6-4 (cit. on p. 54).

[19] M. G. Rojo et al. “Critical Comparison of 31 Commercially Available Digital Slide Systems in Pathology”. In: *International Journal of Surgical Pathology* 14.4 (2006), pp. 285–305. DOI: 10.1177/1066896906292274 (cit. on p. 54).

[20] H. Netten et al. “A fast scanner for fluorescence microscopy using a 2-D CCD and time delayed integration”. In: *Bioimaging* 2.4 (1994), pp. 184–192. DOI: 10.1002/1361-6374(199412)2:4%3C184::AID-BIO3%3E3.0.CO;2-M (cit. on pp. 54, 57).

[21] M. Montalto, R. Filkins, and R. McKay. “Autofocus methods of whole slide imaging systems and the introduction of a second-generation independent dual sensor scanning method”. In: *Journal of Pathology Informatics* 2.1 (2011), p. 44. DOI: 10.4103/2153-3539.86282 (cit. on p. 54).

[22] M. E. Bravo-Zanoguera et al. “Dynamic autofocus for continuous-scanning time-delay-and-integration image acquisition in automated microscopy”. In: *Journal of Biomedical Optics* 12.3 (2007), p. 034011. DOI: 10.1117/1.2743078 (cit. on p. 54).

[23] B. Hulsken. “Autofocus based on differential measurements”. In: *US Patent No. 9,832,365 B2* (2017) (cit. on p. 54).

[24] B. Hulsken. “Method for simultaneous capture of image data at multiple depths of a sample”. In: *US Patent No. 9,910,258 B2* (2018) (cit. on pp. 54, 55, 75).

[25] B. Hulsken. “Scanning imaging system with a novel imaging sensor with gaps for electronic circuitry”. In: *US Patent. No. 10,091,445 B2* (2018) (cit. on pp. 54, 55, 75).

[26] B. Hulsken and S. Stallinga. “Sensor for microscopy”. In: *US Patent No. 10,353,190 B2* (2019) (cit. on pp. 54, 55, 75).

[27] L. van der Graaff et al. “Multi-line fluorescence scanning microscope for multi-focal imaging with unlimited field of view”. In: *Biomedical Optics Express* 10.12 (2019), pp. 6313–6339. DOI: 10.1364/BOE.10.006313 (cit. on pp. 55, 71, 72).

[28] S. M. Shakeri, L. J. van Vliet, and S. Stallinga. “Impact of partial coherence on the apparent optical transfer function derived from the response to amplitude edges”. In: *Applied Optics* 56.12 (2017), p. 3518. DOI: 10.1364/AO.56.003518 (cit. on pp. 55, 56).

[29] T. C. Cornish, R. E. Swapp, and K. J. Kaplan. “Whole-slide Imaging: Routine Pathologic Diagnosis”. In: *Advances In Anatomic Pathology* 19.3 (2012), pp. 152–159. DOI: 10.1097/PAP.0b013e318253459e (cit. on p. 55).

[30] N. Roberts et al. “Toward Routine Use of 3D Histopathology as a Research Tool”. In: *The American Journal of Pathology* 180.5 (2012), pp. 1835–1842. DOI: 10.1016/j.ajpath.2012.01.033 (cit. on p. 55).

[31] E. A. El-Gabry, A. V. Parwani, and L. Pantanowitz. “Whole-slide imaging: widening the scope of cytopathology”. In: *Diagnostic Histopathology* 20.12 (2014), pp. 456–461. DOI: 10.1016/j.mpdhp.2014.10.006 (cit. on p. 55).

[32] W. E. Khalbuss, L. Pantanowitz, and A. V. Parwani. “Digital Imaging in Cytopathology”. In: *Pathology Research International* 2011 (2011), pp. 1–10. DOI: 10.4061/2011/264683 (cit. on p. 55).

[33] T. Kalinski et al. “Virtual 3D Microscopy Using Multiplane Whole Slide Images in Diagnostic Pathology”. In: *American Journal of Clinical Pathology* 130.2 (2008), pp. 259–264. DOI: 10.1309/QAM22Y85QCV5JM47 (cit. on p. 55).

[34] Y. Park, C. Depeursinge, and G. Popescu. “Quantitative phase imaging in biomedicine”. In: *Nature Photonics* 12.10 (2018), pp. 578–589. DOI: 10.1038/s41566-018-0253-x (cit. on p. 55).

[35] M. R. Teague. “Deterministic phase retrieval: a Green’s function solution”. In: *Journal of the Optical Society of America* 73.11 (1983), pp. 1434–1441. DOI: 10.1364/JOSA.73.001434 (cit. on pp. 55, 68).

[36] D. Paganin and K. A. Nugent. “Noninterferometric Phase Imaging with Partially Coherent Light”. In: *Physical Review Letters* 80.12 (1998), pp. 2586–2589. DOI: 10.1103/PhysRevLett.80.2586 (cit. on pp. 55, 68).

[37] L. Waller et al. “Phase and amplitude imaging from noisy images by Kalman filtering”. In: *Optics Express* 19.3 (2011), pp. 2805–2815. DOI: 10.1364/OE.19.002805 (cit. on pp. 55, 68, 76).

[38] E. Bostan et al. “Variational Phase Imaging Using the Transport-of-Intensity Equation”. In: *IEEE Transactions on Image Processing* 25.2 (2016), pp. 807–817. DOI: 10.1109/TIP.2015.2509249 (cit. on pp. 55, 68, 76).

[39] T. Kim et al. “White-light diffraction tomography of unlabelled live cells”. In: *Nature Photonics* 8.3 (2014), pp. 256–263. DOI: 10.1038/nphoton.2013.350 (cit. on p. 55).

[40] R. Horstmeyer et al. “Diffraction tomography with Fourier ptychography”. In: *Optica* 3.8 (2016), pp. 827–835. DOI: 10.1364/OPTICA.3.000827 (cit. on p. 55).

[41] A. Descloux et al. “Combined multi-plane phase retrieval and super-resolution optical fluctuation imaging for 4D cell microscopy”. In: *Nature Photonics* 12.3 (2018), pp. 165–172. DOI: 10.1038/s41566-018-0109-4 (cit. on pp. 55, 65, 69, 76).

[42] S. Abeytunge et al. “Rapid confocal imaging of large areas of excised tissue with strip mosaicing”. In: *Journal of Biomedical Optics* 16.5 (2011), p. 050504. DOI: 10.1117/1.3582335 (cit. on p. 57).

[43] S. M. Shakeri et al. “Optical quality assessment of whole slide imaging systems for digital pathology”. In: *Optics Express* 23.2 (2015), pp. 1319–1336. DOI: 10.1364/OE.23.001319 (cit. on p. 57).

[44] S. Stallinga. “Finite conjugate spherical aberration compensation in high numerical-aperture optical disc readout”. In: *Applied Optics* 44.34 (2005), pp. 7307–7312. DOI: 10.1364/AO.44.007307 (cit. on pp. 57, 58).

[45] G. M. P. van Kempen et al. “A quantitative comparison of image restoration methods for confocal microscopy”. In: *Journal of Microscopy* 185.3 (1997), pp. 354–365. DOI: 10.1046/j.1365-2818.1997.d01-629.x (cit. on p. 62).

[46] R. Heintzmann. “Estimating missing information by maximum likelihood deconvolution”. In: *Micron* 38.2 (2007), pp. 136–144. DOI: 10.1016/j.micron.2006.07.009 (cit. on p. 62).

[47] J. B. Sibarita. “Deconvolution microscopy”. In: *Advances in Biochemical Engineering/Biotechnology* 95 (2005), pp. 201–243. DOI: 10.1007/b102215 (cit. on p. 62).

[48] P. J. Verveer and T. M. Jovin. “Acceleration of the ICTM image restoration algorithm”. In: *Journal of Microscopy* 188.3 (1997), pp. 191–195. DOI: 10.1046/j.1365-2818.1997.2810827.x (cit. on pp. 62, 63).

[49] A. Azaripour et al. “A survey of clearing techniques for 3D imaging of tissues with special reference to connective tissue”. In: *Progress in Histochemistry and Cytochemistry* 51.2 (2016), pp. 9–23. DOI: 10.1016/j.proghi.2016.04.001 (cit. on p. 64).

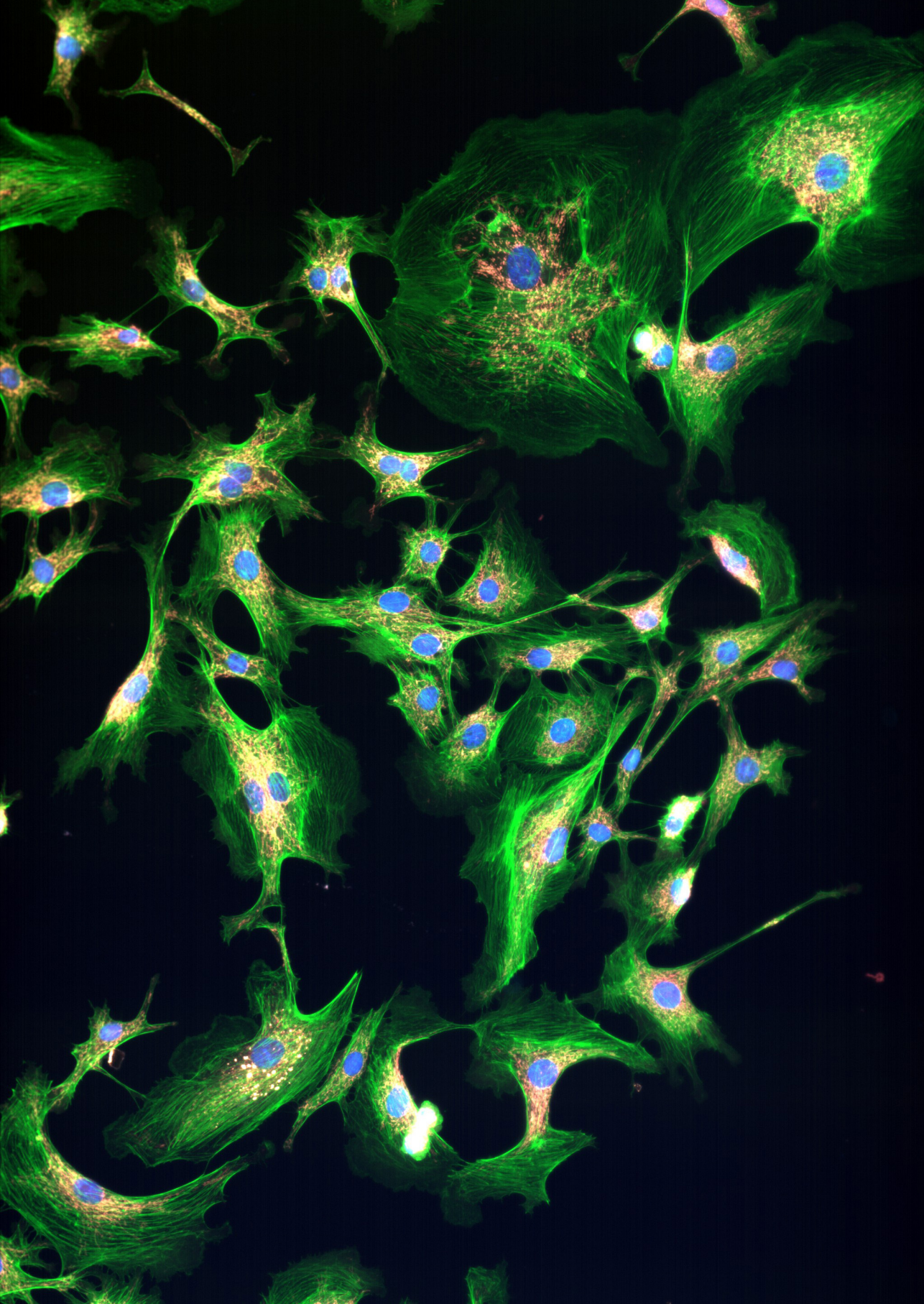
[50] G. Pauletti et al. “Detection and quantitation of HER-2/neu gene amplification in human breast cancer archival material using fluorescence in situ hybridization.” In: *Oncogene* 13.1 (1996), pp. 63–72. URL: http://www.ncbi.nlm.nih.gov/pubmed/8700555 (cit. on p. 71).

[51] L. van der Graaff et al. “Fluorescence imaging for whole slide scanning using LED-based color sequential illumination”. In: *Proc. SPIE 10679, 106790D* (2018). DOI: 10.1117/12.2306776 (cit. on pp. 73, 75).

[52] F. Crameri. “Geodynamic diagnostics, scientific visualisation and StagLab 3.0”. In: *Geoscientific Model Development* 11.6 (2018), pp. 2541–2562. DOI: 10.5194/gmd-11-2541-2018 (cit. on p. 71).

[53] P. Moulin. “Multiscale Image Decompositions and Wavelets”. In: *The Essential Guide to Image Processing*. Elsevier, 2009, pp. 123–142. DOI: 10.1016/B978-0-12-374457-9.00006-8 (cit. on p. 76).

[54] C. Vonesch and M. Unser. “A Fast Thresholded Landweber Algorithm for Wavelet-Regularized Multidimensional Deconvolution”. In: *IEEE Transactions on Image Processing* 17.4 (2008), pp. 539–549. DOI: 10.1109/TIP.2008.917103 (cit. on p. 76).



4

Fluorescence imaging for Whole Slide Scanning using LED-Based Color Sequential Illumination

In the field of pathology there is an ongoing transition to the use of Whole Slide Imaging (WSI) systems which scan tissue slides at intermediate resolution (~0.25 µm) and high throughput (15 mm^2/min) to digital image files. Most scanners currently on the market are line-sensor based push-broom scanners for three-color (RGB) brightfield imaging. Adding the ability of fluorescence imaging opens up a wide range of possibilities to the field, in particular the use of specific molecular (proteins, genes) imaging techniques. We propose an extension to fluorescence imaging for a highly efficient WSI systems based on a line scanning technique using multi-color LED epi-illumination. The use of multi-band dichroics eliminates the need for filter wheels or any other moving parts in the system, the use of color sequential illumination with LEDs enables imaging of multiple color channels with a single sensor. Our approach offers a solution to fluorescence WSI systems that is technologically robust and cost- effective. We present design details of a four-color LED based epi-illumination with a quad-band dichroic filter optimized for LEDs. We provide a thorough analysis regarding the obtained optical and spectral efficiency. The primary throughput limitation is the minimum Signal-to-Noise-Ratio (SNR) given the available optical power in the illumination étendue, and indicates that a throughput on the order of 1000 lines/sec can be obtained.

This chapter has been published as: Leon van der Graaff, Geert J.L.H. van Leenders, Fanny Boyaval, and Sjoerd Stallinga in *Proc. SPIE* **10679** 106790D, (2018) [1].

4.1. Introduction

Interpreting images of tissues and cells for diagnosis is a core activity of pathologists. For over a century, microscopes have been used to visualize these small structures. This has changed, however, with the emergence of digital pathology, which uses digital slide scanners to acquire high resolution (~0.25 µm) digital images of complete tissue slides (~15x15 mm) in a very short time (~1 min). These images, referred to as 'Whole Slide Images' (WSI) or 'virtual slides' are then assessed for diagnosis on a computer screen. Advantages of Digital Pathology include collaboration at a distance, correlation with radiologic images, workflow management and control, teaching, certification, and it opens the way for Computer Aided Diagnostics (CAD) and Clinical Decision Support (CDS)[2–4].

It is a major opportunity for the field of pathology for slide scanners to become compatible with fluorescence imaging. This would open the Digital Pathology environment to the benefits of both immunofluorescence studies (visualizing specific proteins) and FISH studies (Fluorescence in situ hybridization, detecting and localizing specific parts of the DNA or RNA). Both techniques offer the benefits for more reproducible and quantitative diagnosis [5–7]. Current slide scanners, however, are usually not capable of fluorescence imaging [8]. The few commercial solutions that are available, most often lack high throughput or cannot resolve multiple fluorescent dyes.

Various optical architectures for slide scanners are available [9], for example the step-and-stitch approach which uses standard wide field acquisitions and merges them together digitally. However, the most favored architecture appears to be continuous scanning with a line sensor ('push broom' scanning), because of the mechanical simplicity and reduced need for stitching [10]. Most often, these systems make use of TDI (Time Delayed Integration) to increase the SNR [11]. Combining this design with time-sequential illumination removes the need of multiple sensors for different color channels and gives intrinsic alignment of the color channels.

Here, we propose an extension of this architecture to fluorescence scanning that is simple and robust. To achieve this, a fast switchable, high power light source is required. For that, LEDs are the most obvious choice. Advantages are a high optical power, low price, long lifetime, small size, and their availability in a wide range of colors. Usually, multi-color fluorescence imaging systems make use of mechanical filter wheels. This is, however, incompatible with line scan rates in excess of kHz rates. In our setup we have chosen to use a single 'quad-band' dichroic filter with four different reflection- and passbands and corresponding emission and excitation filters.

An intrinsic limitation to high throughput wide-field fluorescence imaging is the relative low light level that is detected, typically several orders of magnitude less than is available in brightfield imaging. This will have a big influence on the tradeoff between exposure times and the shot-noise dominated SNR, i.e. scanning will be either slow or noisy. This article will therefore focus on the efficiency of the illumination and its effect on the image quality. In the methods section, the spectral, optical and electrical efficiencies will be discussed. In our results, we describe the obtained illumination power and the resulting scanning speed that can be obtained given the required SNR.

4.2. Methods

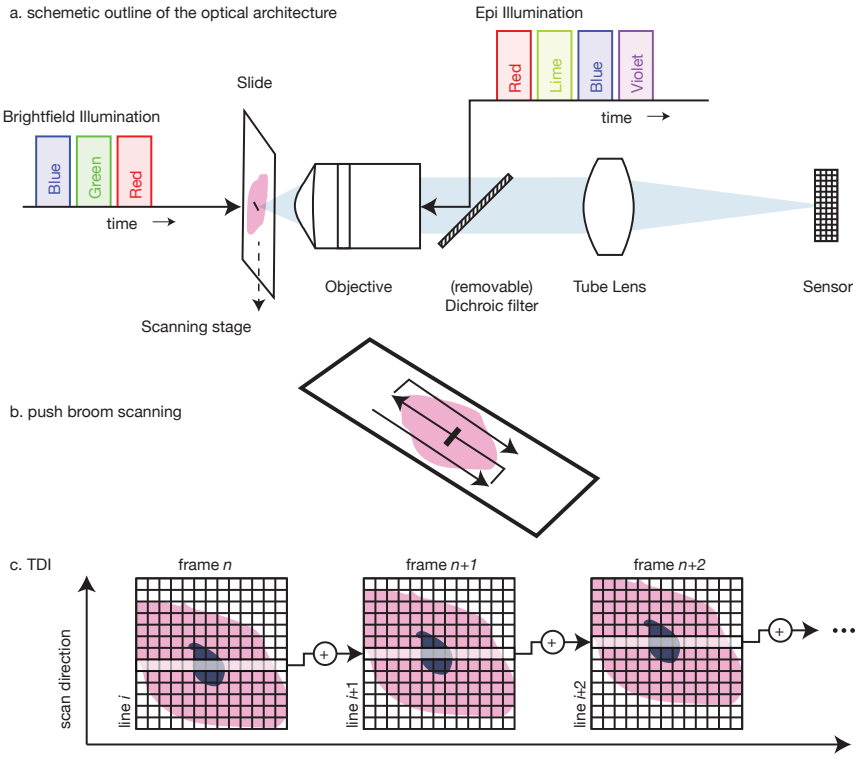


Figure 4.1: Illustration of the scanning system.
(a) A microscope slide is translated by a scanning stage and imaged line by line on a monochromatic sensor. It is illuminated using either brightfield illumination or epi-illumination. Brightfield illumination is done sequentially with red green and blue to obtain an RGB image. For epi-illumination, the dichroic mirror is placed in the optical path and the sample is illuminated sequentially with red, lime, blue and violet to obtain multi-channel fluorescence images.
(b) 'Push-broom scanning': the slide is scanned in arbitrary long lanes, back and forth over the slide.
(c) Time Delayed Integration (TDI): when using a multi-line sensor, frames of sequential acquisitions can be combined to increase the signal level. For this, the movement of the slide has to be synchronized to the readout.

4.2.1. Scanning System

Our whole slide scanner prototype is based on the architecture described by Shakeri et al. [10], see Fig. 4.1. We use a Nikon 20× NA 0.75 Plan Apochromat VC objective lens and a custom Nikon tube lens (effective focal length of 222.4 ± 2.2 mm) for obtaining a target magnification of $M = 22.24$. For lower magnification images, a Nikon 10× NA .45 Plan Apochromat λ objective is used. A sample stage is made using two stages: a PI M-505 low profile translation stage is used for positioning of the slide in the field direction and a Newport XM1000 ultra precision linear motor stage for the continues scanning motion of the sample. To focus, we used two stages to axially translate the objective: a PI M-111 compact micro-translation stage for coarse positioning, and a PI P-721.CL0 piezo nano-positioner

for fine positioning. Images are captured using a Hamamatsu Orca Flash 4.0 v2 CMOS Camera. This sensor has a pixel pitch of 6.3 µm corresponding to 0.29 µm in object space. The illumination is controlled by a National Instruments NI PXIe-6363 data acquisition card. The camera and illumination control are synchronized to the trigger output of the scanning stage.

4.2.2. Spectral efficiency

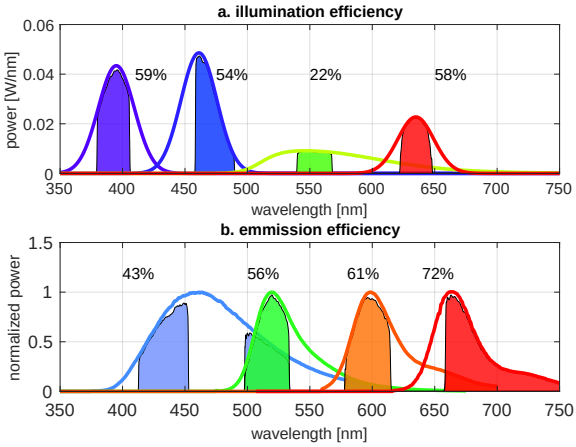


Figure 4.2: Spectral efficiency of the quad-band filter set.
(a) Power spectrum of the LEDs with the power spectrum after filtering indicated by the shaded area. LED and filter spectra are obtained from the data sheets provided by the manufacturers.
(b) Normalized power spectrum of DAPI (blue) Alexa Fluor 488 (green) Mitotracker Red (orange) and cy5 (red) with shaded the power after filtering. Fluorophore spectra are obtained from http://spectra.arizona.edu

In our system we use a quad-band dichroic filter to couple in the epi-illumination. This provides us with four fixed color bands that are matched to the LED sources. This approach is compatible with color sequential illumination. In contrast to using a filter-wheel, it requires no switching time, nor synchronization and has no moving parts. Although Acousto optical tunable filters (AOTF) could provide short switching times and great spectral flexibility, they lack efficiency due their polarization sensitivity and cannot provide the suppression of excitation light needed for fluorescence applications. Using four color bands gives a good coverage of the visible spectrum and provides a good trade-off between the number of colors and spectral efficiency, as will be shown in the next paragraph.

The dichroic filter set in our setup is a Semrock BrightLine Pinkel filter set, optimized for LEDs (LED-DA/FI/TR/Cy5-4X-A-000). The used LED sources are a Ledengin LZ1 Violet LED (~393 nm, LZ1-10UB00-00U5), and the Lumiled Luxeon Rebel Blue (~470 nm, LXML-PB02), Lime (broad spectrum around typical wavelength 554 nm, LXML-PX02-0000) and Red (~635 nm, LXM5-PD01) LEDs. These LEDs are manufactured as a surface mount device (SMD) and have an integrated lens of glass (Ledengin) or silicone (Lumiled) on top of the actual light emitting chip. The LEDs are soldered on metal-core printed circuit boards (MCPCB) for heat dissipation. The calculated spectral efficiencies of this configuration are shown in Fig. 4.2. The top row shows the efficiency of illumination, i.e. the product of the LED spectrum and the filter emission band, which is in general over 50 %. The lime LED has a considerable lower efficiency due to its broad spectrum.

However, lime LEDs have a very high output power compared to green LEDs. This combination appears to provide more power than using an intrinsically less efficient green LED. The bottom row of Fig. 4.2 shows the match of four commonly used fluorophores to the filter set. Again, a typical efficiency of about 50 % is expected.

4.2.3. Étendue efficiency

LEDs are available in a wide range of sizes and powers. The maximum power, however, that can be transmitted through the objective and projected onto the sample appears not to be related to the power of the light source P_s, but rather to its emission intensity P_s/A_s, where A_s is the emission area of the LED. The reason for this is that in a lossless optical system, the étendue (product of area and the solid angle the source subtends) is conserved. In our case, the effectively used circular area of the LED surface $A_s = \pi R_s^2$ and the opening angle α used to collect the light are therefore limited by the étendue E of the objective:

$$\pi R_s^2 \pi n_s^2 \sin^2 \alpha \leq E = \pi \mathrm{NA}^2 \pi \mathrm{FOV}^2 \tag{4.1}$$

where n_s is the refractive index of the source material, NA the numerical aperture of the objective and FOV the radius of the used field of view. Two conclusions can be drawn from this inequality. First, for obtaining optimal efficiency, the LED should have a surface area $\pi R_s^2 \geq E/\pi n_s$, otherwise the étendue of the objective cannot be filled. Second, an estimate can be made of the étendue efficiency by considering the LED to be a Lambertian emitter. The power that can be transmitted through the objective is then given by

$$P = \frac{P_s}{A_s} \pi R_s^2 \sin^2 \alpha \leq P_s \frac{E}{A_s n_2^2 \pi} \equiv P_s \eta_E \tag{4.2}$$

where η_E is the étendue efficiency and where Equation 4.1 is used. This shows that LEDs can be chosen arbitrary large, but only an increase of the surface intensity P_s/A_s will increase the transmitted power. Filling in the relevant parameters for our system (NA = 0.75, FOV = 0.5 mm, n_s = 1.5, A_s = 1 mm^2) gives an estimated maximum obtainable étendue efficiency η_E = 19 %.

4.2.4. Optical design

The optical system is designed to have an even illumination distribution. This is achieved by using a 4f-system to image the LED sources in the back focal plane of the objective. The illumination distribution in the front focal plane (the sample plane) therefore follows the radial emission pattern of the LEDs, which is smooth and fairly uniformly distributed. An even flatter distribution profile can be obtained by having a diffuser in the back focal plane of the collector lenses at the expense of a reduction in illumination power. Therefore, a flat-field correction in post-processing is preferred. The 4f-system has a magnification of 7.5, which is made with an f = 150 mm field lens (Thorlabs AC254-150-A) and f = 20 mm collector lenses (Thorlabs ACL2520U-A) for the LEDs. A schematic layout of the optical design is shown in Fig. 4.3.

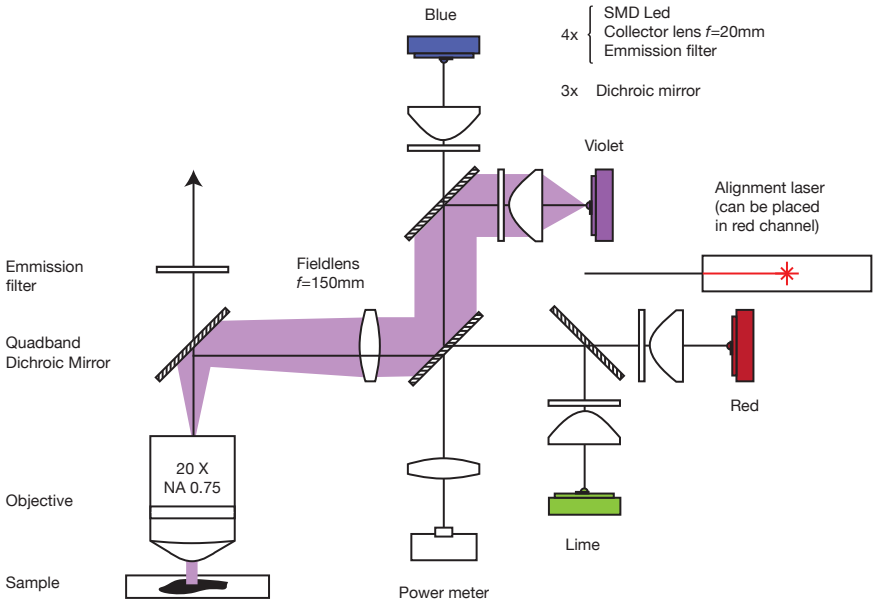


Figure 4.3: Schematic layout of the optical design of the LED-based epi-illumination. The optical path of the violet channel is shown shaded.

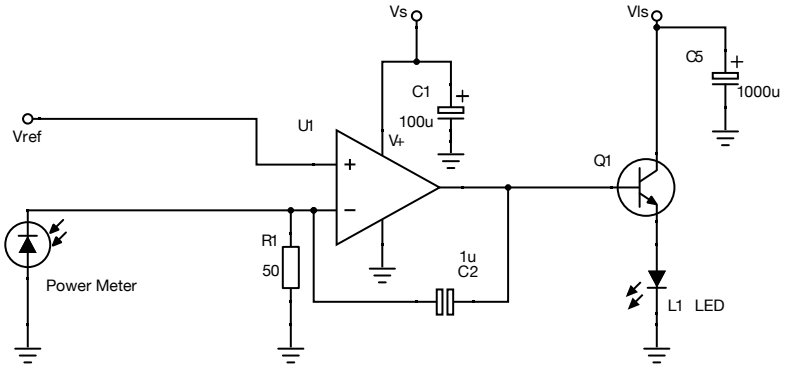


Figure 4.4: Schematic layout of the driver electronics. The optical feedback stabilizes the output of the LED (L1) such that the voltage measured by the power meter is equal to V_{ref}. The system response-time is set by R1C2, which has to be matched to the speed of the power meter in order to prevent instabilities. C1 and C5 are used for stabilization of the power supply.

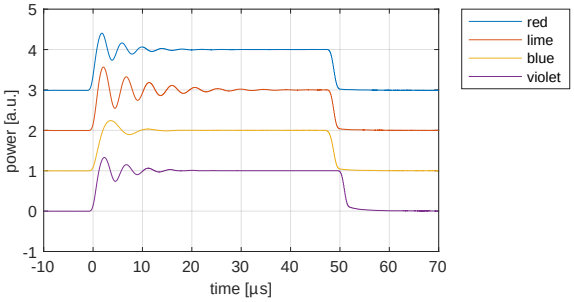


Figure 4.5: Switching times of the driver electronics for all four LED channels. Shown is a 50 µs pulse. The typical rise time is 2 µs. Some channels show ringing effects, which last at most 40 µs after switch-on.

4.2.5. Driver electronics

LEDs are current driven devices. In most applications, this current is either provided by the combination of a voltage source and a series resistor, or by a switched-mode power supply. Both methods are, however, not suitable for color sequential illumination, which requires short switching times (~µs) and high thermal stability. Moreover, the LEDs are turned on only a short and limited period of time (in our case, a few ms and at most a 25 % duty cycle), which gives room for using higher currents than the maximum specified by the manufacturer. Therefore, a fast switching, high current LED driver is developed. This device makes use of optical feedback such that the current can be adjusted to have a constant output power. The feedback compensates fluctuations in LED efficiency due to temperature changes. This removes the need of warming up the LEDs to a steady state temperature, and most importantly, it makes the system insensitive to all settings influencing the average current (e.g. scan speed, pulse length, number of channels) and environmental influences (e.g. room temperature, drift). The output power is measured by a Thorlabs photodetector (PDA36A) that is incorporated in the beam combiner, see Fig. 4.3. Sharing a single detector between the color channels is possible because the LEDs are only used individually. The LED driver makes it possible to drive the LEDs with currents up to 2 A and provides switching times of at most 2 µs while it needs at most 40 µs settling time, see Figs. 4.4 and 4.5. The speed of the driver is currently limited by the bandwidth of the photodetector.

4.2.6. Test slides

Two slides are used to demonstrate the system. The fist slide is a Thermofisher F36924 FluoCells Prepared Slide #1. This slide contains bovine pulmonary artery endothelial (BPAE) cells stained with a combination of fluorescent dyes. Mitochondria are labeled with red-fluorescent MitoTracker Red CMXRos which can be excited using the lime channel, F-actin is stained using green-fluorescent Alexa Fluor 488 phalloidin which can be excited using the blue channel, and blue-fluorescent DAPI which can be excited with the violet channel, labels the nuclei.

The second slide is an immunofluorescence staining of a stage 3 human rectum cancer. The slide is stained with three antibodies. Desmin (IgG1 M Alexa Fluor 488) is highly expressed in muscle cells and can be excited using the blue channel. CD31 (IgG R Alexa Fluor 546) is a marker for blood vessels. It is strongly expressed in endothelial cells and is therefore used to visualize the vasculature in normal and

Table 4.1: Measured optical throughput of the illumination.

Channel	Spectral efficiency	Étendue & Transfer efficiency	Output power mW
Violet	77%	20%	179
Blue	78%	9%	73
Lime	21%	10%	33
Red	72%	11%	82

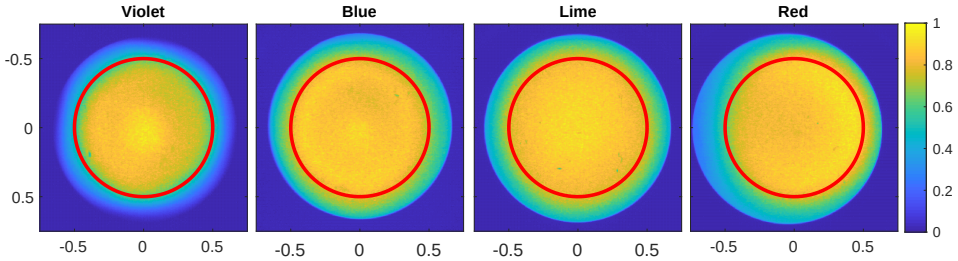


Figure 4.6: Measured distribution of the epi-illumination in the front focal plane of the objective. The red circle has a diameter of 1 mm and indicates the region that is actually used for imaging.

pathological conditions. This stain can be excited using the lime channel. D2-40 (direct labelled with Alexa Fluor 594) is used as a marker of lymphatic endothelium. Although this stain cannot be excited at the most optimal wavelength using our setup, it does show significant fluorescence under violet excitation. Additionally, the slide is stained with DAPI, labelling the nuclei.

4.3. Results

4.3.1. Optical performance

The measured optical output of the designed epi-illumination unit is given in Table 4.1. The spectral efficiency is determined by measuring the LED source power both with and without the excitation filter applied. The measured efficiencies are in general somewhat higher than the expected numbers based on the theoretical spectra, listed in Fig. 4.2. The étendue & transfer efficiency is determined by measuring the fraction of the source power that passes a circular area with a diameter of 1 mm in the front focal plane of the objective, i.e. the fraction that can actually be used for imaging. This number includes both the étendue efficiency, and the losses due to scattering and absorption. The found numbers compare reasonably well to the derived theoretical value of 19 % for a Lambertian emitter. The blue, lime and red channel loose about 50 %. The violet channel performs on par with the theoretical expectation. An explanation for the fact that the results of this channel differs significantly from the others might be found in the different LED optic. Additionally, the light of this LED might be more collimated than the assumed Lambertian emitter. The last column of Table 4.1 shows the optical power that can be effectively used during a 1 kHz, 25 % duty cycle pulse of 2 A peak current.

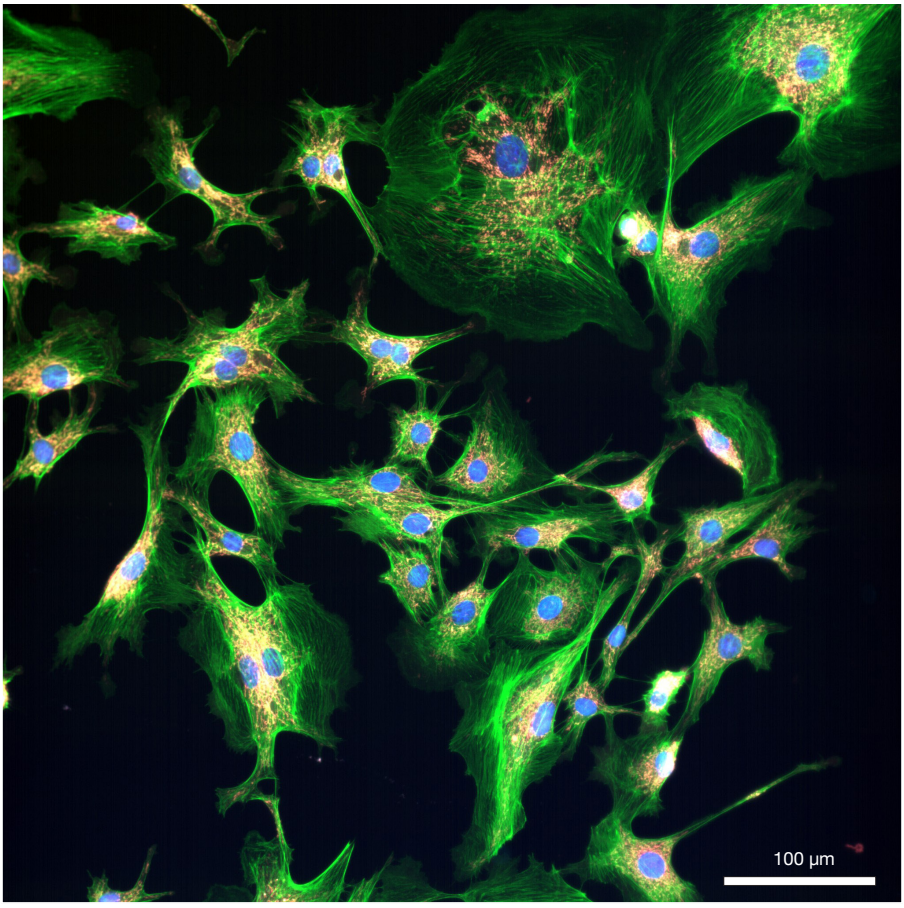


Figure 4.7: Scan of the FluoCells slide using 1 ms exposure time and 96 TDI lines. The displayed image is in false colors: blue corresponds to the violet excitation and labels the nuclei, green to the blue excitation and labels F-actin and red to the lime excitation and labels the mitochondria.

Figure 4.6 shows the distribution of the light in the front focal plane of the objective, which is measured by a Thorlabs CMOS camera (DCC1645C). The distribution appears rather smooth and the power remains over 70% of the peak power throughout the whole region of interest (indicated by the red circle with a diameter of 1 mm. The red channel shows some asymmetry which is caused by a slight misalignment of the LED on the PCB.

4.3.2. Scan results

Three scans are made to demonstrate the capabilities of the system. The first is a scan of the FluoCells slide, shown in Fig. 4.7. The image shows a single-lane scan of 2048×2048 pixels (594 µm×594 µm). The scan is made using the 20× NA .75 objective lens and an exposure time of 1 ms. TDI is implemented by selecting a 96 line region of interest on the camera. The lines are aligned and summed in post-processing. There is no correction done for background level, flat-field or

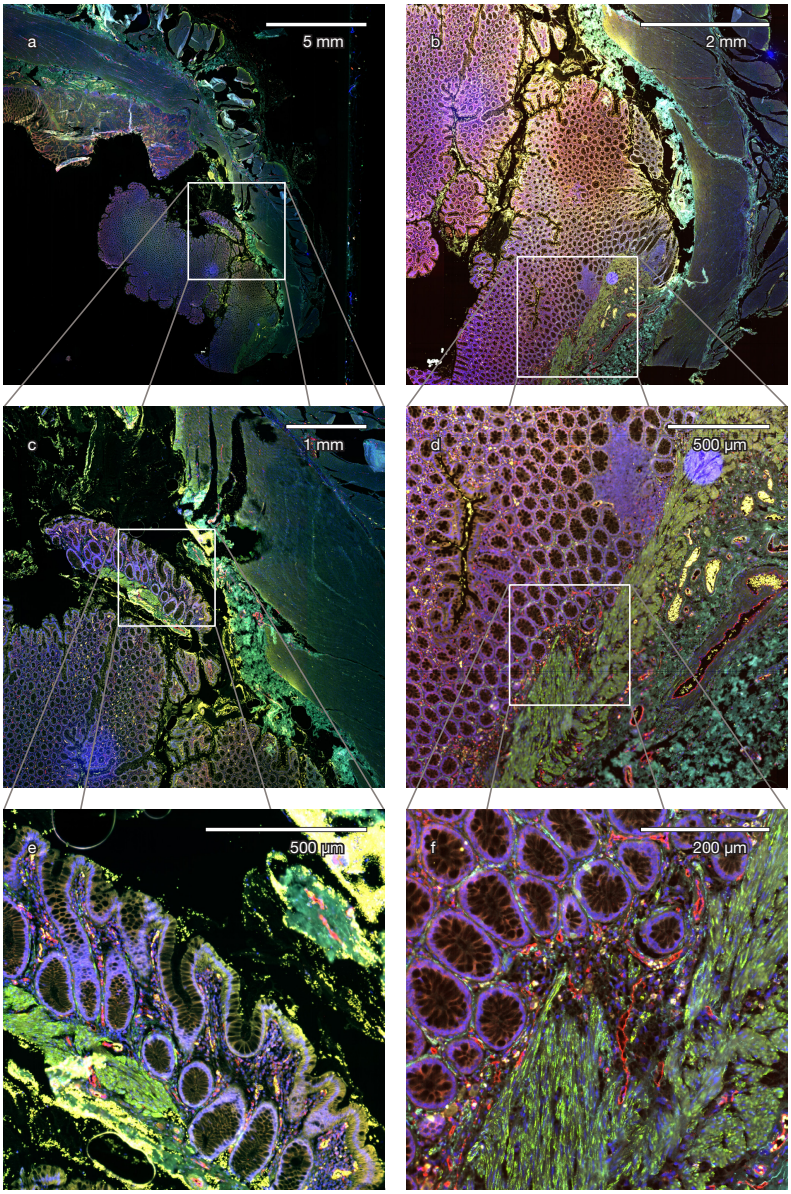


Figure 4.8: Scan of the human rectum slide at different zoom levels.
(a,c,e) Acquired using the 10× NA .45 objective, an exposure time of 1 ms and 96 TDI lines.
(b,d,f) Acquired using the 20× NA .75 objective, an exposure time of 1 ms and 16 TDI lines.

crosstalk. The image shows that a high quality scan can be obtained with an exposure time of 1 ms.

The second scan is given in Figs. 4.8a, 4.8c and 4.8d. The images show a whole slide scan of the human rectum tissue slide at different zoom levels. Figure 4.8a shows a whole slide scan constructed by stitching 16 scan lanes together, resulting in a 32,768×40,000 pixel image (~19×23 mm). Figure 4.8c shows an intermediate zoom level and Figure 4.8d has the width of a single scan lane (~1.2 mm). The scan is made using the 10× NA .45 objective lens. An exposure time of 1 ms is used while the violet channel is reduced to 17 % output power to avoid overexposure. A region of 96 lines on the sensor is used for TDI in post-processing. A correction is applied for background level and flat-field.

The third scan, given in Figs. 4.8b, 4.8d and 4.8f shows the same slide but scanned with the 20× NA .75 objective lens. Figure 4.8b shows 10 scan lanes stitched together, resulting in a 20,480×24,000 pixel image (~6×7 mm). Figure 4.8d shows an intermediate zoom level and Figure 4.8f has the width of a single scan lane (~0.6 mm). The same exposure time (1 ms) and illumination settings where used, but for this scan, only 16 TDI lines where used. A correction is done for background level and flat-field.

4.3.3. Photo-electron yield

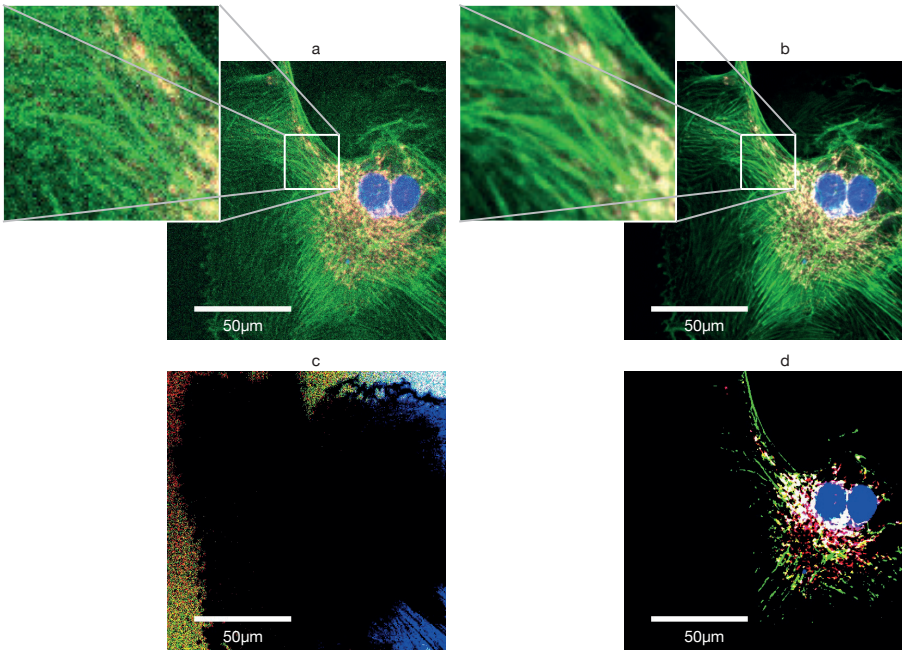


Figure 4.9: Thermofisher F36924 FluoCells Prepared Slide #1 in false colors. The images are made in a static alignment.
(a) Image made using an exposure time of 1 ms.
(b) Image made using an exposure time of 50 ms.
(c) Background area, given by the 5 % pixels with lowest intensities in (b).
(d) Representative area, given by the 5 % pixels with highest intensities in (b).

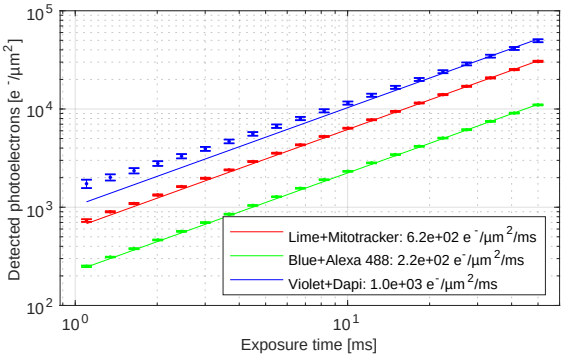


Figure 4.10: Photo-electron counts per µm^2 (measured in object space) as a function of exposure time for the 3 stains of the FluoCells slide. Error bars indicate 2 standard deviations. The straight lines indicate a linear fit with the exposure time, of which the numerical results are shown in the legend. The conversion 1 e$^-$/px = 11.7 e$^-$/µm^2 is used based on the pixel pitch of 0.29 µm.

The SNR of an optical system is ultimately limited by shot noise, which has a $\sqrt{N}$ standard deviation for N collected photo-electrons. It is therefore relevant to test how many photo-electrons can be captured in a practical application. To this end, a series of images is acquired of the FluoCells slide in a static alignment (i.e. without scanning) to allow repeating the experiment multiple times without alignment errors. A range of exposure times between 1 and 50 ms is used in order to vary the illumination dose. An example of an image acquired with an exposure time of 1 ms is given in Fig. 4.9a and with an exposure time of 50 ms in Fig. 4.9b. A uniform background level is present in the images, mainly due to autofluorescence. This level is determined by averaging over an area of the image where no cell structure is present, defined by selecting the 5 % pixels with the lowest intensity in the image acquired using an exposure time of 50 ms, see Fig. 4.9c. Subsequently, the background for each channel is subtracted. The representative signal is determined from averaging an area of the image with high intensities, defined by selecting the top 5 % pixels with the highest intensity in the image acquired using an exposure time of 50 ms, see Fig. 4.9d.

For every exposure time, ten noise independent images are acquired and processed. Fig. 4.10 shows the obtained average number of collected photo-electrons per µm^2 (measured in object space) as a function of the exposure time. The violet channel performs very well, which can be explained by the high power of the LED and the strong fluorescence of the DAPI stain. In contrast, the number of photons obtained in the blue channel is quite low, even considering the relative power of the illumination.

The part of the human rectum slide shown in Fig. 4.8f is analysed correspondingly. In this patch, the blue and lime channel have a photon yield of 1.4×10^2 and 9.0×10^2 e$^-$/ms/µm^2, which is on the same order of magnitude as the yield of the FluoCells slide. The violet channel has a yield of 2.0×10^4 e$^-$/ms/µm^2, where is corrected for the reduced excitation power. This is significantly higher than the other stains.

4.3.4. FRC Resolution

The impact of shot noise on image quality can also be assessed by its effect on the information content of the images. This is done by determining the Fourier

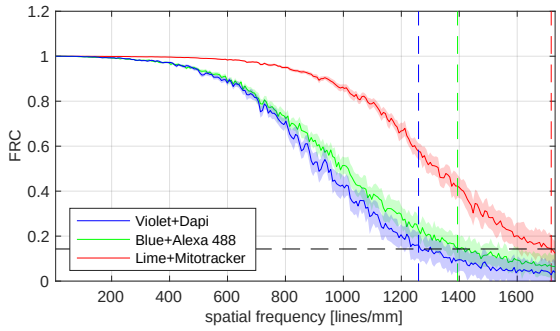


Figure 4.11: FRC curves of the FluoCells slide for an exposure time of 10 ms. The shaded area indicate ± 2 standard deviations. The colored dashed lines indicate the FRC resolution. The black dashed line is the noise threshold of $1/7^{\text{th}}$.

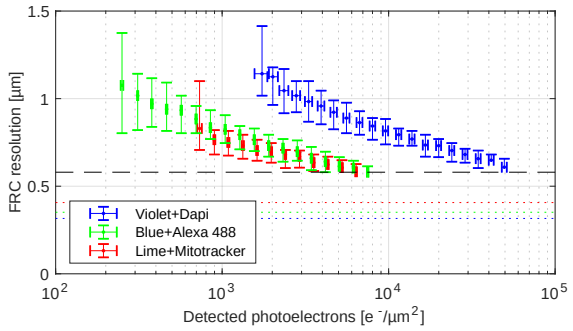


Figure 4.12: FRC resolution of the FluoCells slide as a function of the photo-electron counts/µm^2. Error bars indicate 2 standard deviations. The dotted colored lines indicate the diffraction limit for the different fluorescent labels. The dashed black line gives the Nyquist sampling limit given the pixel size and magnification of the system.

Ring Correlation (FRC) resolution [12–14] of the images as a function of the photo-electron count. The FRC resolution method determines the spatial frequency up to which the correlation between noise independent image acquisitions is higher than a noise threshold (typically, FRC= 1/7 is used). The inverse of this spatial frequency is the FRC resolution, and measures the smallest detail that can reliably be discerned in the image. The advantage of the FRC method is that it includes all effects of the optical imaging system, such as the noise level and the optical transfer function, but also the underlying sample structure itself. Here, we anticipate that a low SNR will lead to a smaller FRC resolution.

Ten noise independent acquisitions where done and the FRC is calculated for every independent pair. This provides an average and a standard deviation of the entire FRC curve, see for example Fig. 4.11. The FRC resolution and the associated uncertainty can then be deduced from the threshold criterion. The obtained results are plotted in Fig. 4.12. The measured values for the blue and lime channel are very similar: the FRC resolution improves significantly by increasing the photo-electron counts up to $10^4 e^-/\mu m^2$, at which the Nyquist sampling distance of the sensor (580 nm) is reached. It can therefore be concluded that for photo-electron counts below this number, the effective FRC resolution is limited by the shot-noise. Linear extrapolation to very high photo-electron counts suggests that about $10^5 e^-/\mu m^2$ have to be collected to approach the diffraction limit at $\lambda/2NA$. The violet channel shows significantly different results, which is attributed to the lack of high frequent content in the underlying sample structure compared to the

blue and lime channel.

4.4. Conclusion & Discussion

A high throughput whole slide scanning system was extended to fluorescence imaging. This was achieved by adding a LED based epi-illumination. We used multi-band dichroics and therefore no moving parts were required in the system. Our system is compatible with color sequential illumination, which enables multi color imaging using a single monochromatic sensor. The light source produces a smooth illumination pattern with rim intensities better than 70 % and has a high and stable output power. In practice, our system can capture at least 200 to 1000 $e^-/\mu m^2/ms$. It was found that about $10^4 e^-/\mu m^2$ are required to obtain an FRC resolution matching the Nyquist sampling length of the sensor (560 nm). Extrapolation of our results suggest that about an order of magnitude more photons are required to obtain true diffraction limited resolution.

Most WSI systems use continues scanning with a line sensor because of the mechanical simplicity and reduced need for stitching. The required exposure time is the main speed limitation in fluorescence imaging with LED illumination. This stands in contrast to brightfield imaging, where the throughput is limited by the digital bandwidth. Intrinsically, a step-and-stitch approach will therefore be faster than (single) line scanning for fluorescence imaging with incoherent LED illumination because it makes better use of the illumination étendue. This disadvantage of line scanning is partly overcome by the use of TDI.

Consider, for example, a three-color fluorescence image with moderate SNR. Our numbers suggest that about $10^4 e^-/\mu m^2$ have be be acquired. For the dye with the lowest yield we measured (Alexa Fluor 488 in the FluoCells slide), this needs 50 ms exposure time. Using a single-line system with a field of view of 1 mm and a pixel size of 0.25 µm, scanning an area of 15 mm × 15 mm in 3 colors would then take 38 h. It would require the use of 96 TDI lines to reduce the scanning time to a reasonable 23 min.

Our proposed method for extending a WSI platform to fluorescence scanning is a technologically robust and cost effective solution while adding just a few components. The obtained scanning speed is sufficient for most clinical applications where occasional scanning of fluorescent slides is needed. In this way, the benefits of fluorescence imaging can become available for the users of Digital Pathology.

References

[1] L. van der Graaff et al. “Fluorescence imaging for whole slide scanning using LED-based color sequential illumination”. In: *Optics, Photonics, and Digital Technologies for Imaging Applications V*. Vol. 10679. International Society for Optics and Photonics. SPIE, 2018. ISBN: 9781510618848. DOI: 10.1117/12.2306776 (cit. on p. 83).

[2] M. Kowal et al. “Computer-aided diagnosis of breast cancer based on fine needle biopsy microscopic images.” In: *Computers in biology and medicine* 17.10 (2011), pp. 1563–1572 (cit. on p. 84).

[3] L. Langer et al. “Computer-aided diagnostics in digital pathology: automated evaluation of early-phase pancreatic cancer in mice”. In: *International Journal of Computer Assisted Radiology and Surgery* 10.7 (2015), pp. 1043–1054. DOI: 10.1007/s11548-014-1122-9 (cit. on p. 84).

[4] C. Castaneda et al. “Clinical decision support systems for improving diagnostic accuracy and achieving precision medicine”. In: *Journal of Clinical Bioinformatics* 5.1 (2015), p. 4. DOI: 10.1186/s13336-015-0019-3 (cit. on p. 84).

[5] F. Ghaznavi et al. “Digital Imaging in Pathology: Whole-Slide Imaging and Beyond”. In: *Annual Review of Pathology: Mechanisms of Disease* 8.1 (2013), pp. 331–359. DOI: 10.1146/annurev-pathol-011811-120902 (cit. on p. 84).

[6] C. Higgins. “Applications and challenges of digital pathology and whole slide imaging”. In: *Biotechnic & Histochemistry* 90.5 (2015), pp. 341–347. DOI: 10.3109/10520295.2015.1044566 (cit. on p. 84).

[7] H.-A. Lehr et al. “Quantitative Evaluation of HER-2/ neu Status in Breast Cancer by Fluorescence In Situ Hybridization and by Immunohistochemistry With Image Analysis”. In: *American Journal of Clinical Pathology* 115.6 (2001), pp. 814–822. DOI: 10.1309/AJ84-50AK-1X1B-1Q4C (cit. on p. 84).

[8] L. Pantanowitz, N. Farahani, and A. Parwani. “Whole slide imaging in pathology: advantages, limitations, and emerging perspectives”. In: *Pathology and Laboratory Medicine International* 7 (2015), p. 23. DOI: 10.2147/PLMI.S59826 (cit. on p. 84).

[9] M. G. Rojo et al. “Critical Comparison of 31 Commercially Available Digital Slide Systems in Pathology”. In: *International Journal of Surgical Pathology* 14.4 (2006), pp. 285–305. DOI: 10.1177/1066896906292274 (cit. on p. 84).

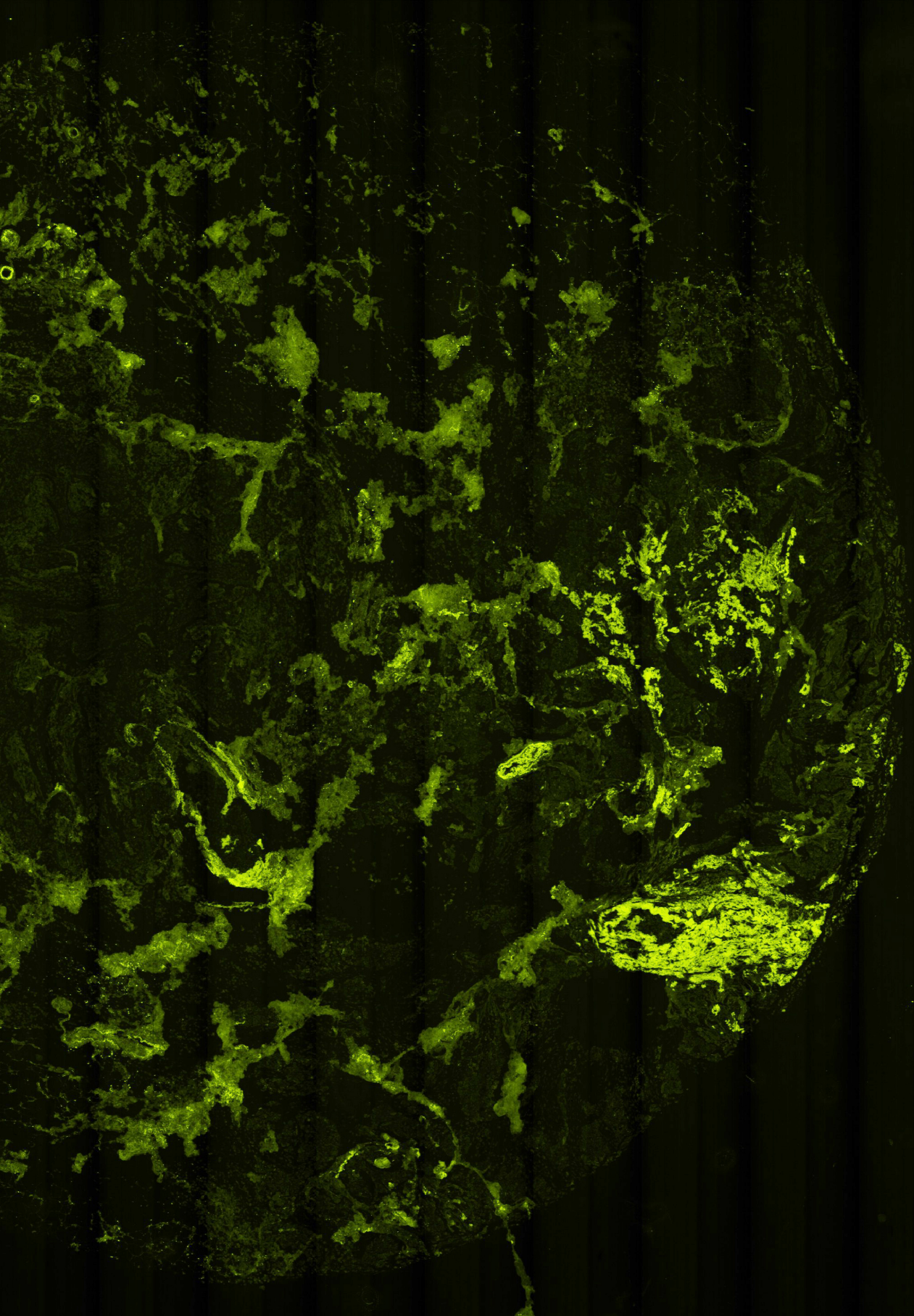
[10] S. M. Shakeri et al. “Optical quality assessment of whole slide imaging systems for digital pathology”. In: *Optics Express* 23.2 (2015), pp. 1319–1336. DOI: 10.1364/OE.23.001319 (cit. on pp. 84, 85).

[11] H. Netten et al. “A fast scanner for fluorescence microscopy using a 2-D CCD and time delayed integration”. In: *Bioimaging* 2.4 (1994), pp. 184–192. DOI: 10.1002/1361-6374(199412)2:4%3C184::AID-BIO3%3E3.0.CO;2-M (cit. on p. 84).

[12] R. P. J. Nieuwenhuizen et al. “Measuring image resolution in optical nanoscopy”. In: *Nature Methods* 10.6 (2013), pp. 557–562. DOI: 10.1038/nmeth.2448 (cit. on p. 95).

[13] P. B. Rosenthal and R. Henderson. “Optimal Determination of Particle Orientation, Absolute Hand, and Contrast Loss in Single-particle Electron Cryomicroscopy”. In: *Journal of Molecular Biology* 333.4 (2003), pp. 721–745. DOI: 10.1016/j.jmb.2003.07.013 (cit. on p. 95).

[14] M. Van Heel. “Similarity measures between images”. In: *Ultramicroscopy* 21.1 (1987), pp. 95–100. DOI: 10.1016/0304-3991(87)90010-6 (cit. on p. 95).



5

Multi-line fluorescence scanning microscope for multi-focal imaging with unlimited field of view

Confocal scanning microscopy is the *de facto* standard modality for fluorescence imaging. Point scanning, however, leads to a limited throughput and makes the technique unsuitable for fast multi-focal scanning over large areas. We propose an architecture for multi-focal fluorescence imaging that is scalable to large area imaging. The design is based on the concept of line scanning with continuous 'push broom' scanning. Instead of a line sensor, we use an area sensor that is tilted with respect to the optical axis to acquire image data from multiple depths inside the sample simultaneously. A multi-line illumination where the lines span a plane conjugate to the tilted sensor is created by means of a diffractive optics design, implemented on a Spatial Light Modulator. In particular, we describe a design that uses higher order astigmatism to generate focal lines of substantially constant peak intensity along the lines. The proposed method is suitable for fast 3D image acquisition with unlimited field of view, it requires no moving components except for the sample scanning stage, and provides intrinsic alignment of the simultaneously scanned focal slices. As proof of concept, we have scanned 9 focal slices simultaneously over an area of 36 mm^2 at 0.29 µm pixel size in object space. The projected ultimate throughput that can be realized with the proposed architecture is in excess of 100 Mpixel/s.

This chapter has been published as: Leon van der Graaff, Geert J.L.H. van Leenders, Fanny Boyaval, and Sjoerd Stallinga in *Biomedical Optics Express* **10**, 12 (2019) [1].

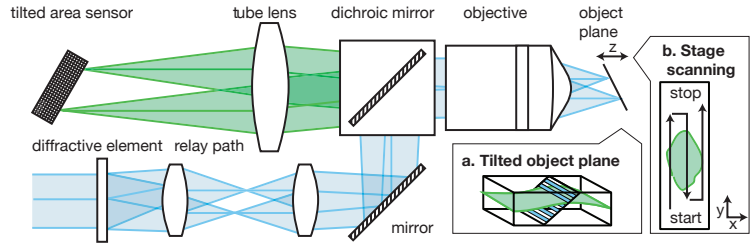


Figure 5.1: Schematic side view of the architecture for a multi-line fluorescence scanning microscope for multi-focal imaging with an unlimited field of view. For clarity we illustrate the concept using only two focal layers. The imaging path is shown in green, the illumination path is shown in blue. A diffractive optical element (DOE) is used for generating a set of parallel scan lines, focused at equidistant planes. Note that in this side view, the lines are orthogonal to the drawing. (a) The tube lens and objective lens form a telecentric optical system such that the tilted sensor has a tilted conjugate plane in object space. The rows on the area sensor conjugate to the line foci are used to capture the image data from the different focal slices scanned by the line foci simultaneously. (b) The sample is scanned in a continuous 'push broom' scanning fashion to obtain a multi-focal image with a field of view that is in principle unlimited.

5.1. Introduction

In the current era of big data analysis, instrumentation to generate massive amounts of image data is in high demand. For high throughput screening in biology, or for novel computer aided medical diagnoses in the field of digital pathology, there is a need for fluorescence imaging of tissues over large fields of view (~few cm), in 3D (up to hundred layers of µm thickness), and at cellular resolution (~1 µm) [2–4]. This can be used, for example, in immunofluorescence or fluorescence *in situ* hybridization (FISH) studies.

The *de facto* standard modality for fluorescence imaging is scanning confocal microscopy [5], because the optical sectioning capability enables high contrast. The underlying point scanning technique has a limited throughput, and the imaged area is limited by the Field Of View (FOV) of the microscope objective. Parallelization is a strategy to increase throughput, as then the space-bandwidth-time product is increased [6]. For example, in spinning disk microscopy, a large number of points is scanned in parallel [7, 8]. Wide field structured illumination has also been proposed as a technique for high throughput imaging [9, 10], where the loss in resolution of lower Numerical Aperture (NA) objectives for increasing the FOV, is compensated by the use of structured illumination. Throughput can also be increased by scanning multiple depth layers in parallel using an illumination with multiple foci [11, 12]. These techniques require the distribution of the emitted light over several detector arms to apply a pinhole conjugate to the foci that scan the specimen at different depths. The necessary beam splitters used in these approaches result in a loss in collected fluorescence signal strength by a factor equal to the number of scanned layers. Another throughput enhancing technique is the use of a line illumination instead of a spot illumination in combination with a line sensor [13–15]. The pinhole for achieving optical sectioning must then be replaced by a slit, which goes at the expense of a small loss in the optical sectioning capability [14, 16]. Line scanning can be combined with multi-focus scanning [17], but the proposed method suffers from the same signal losses induced by splitting the

beam in the detection path as the point scanning based multi-focus systems.

Scanning large, cm^2 sized, areas with all mentioned systems can be accomplished by using 'mosaic' or 'step-and-stitch' scanning. The most favorable method for scanning such large areas, however, is continuous 'push broom' scanning with a line sensor, because of its mechanical simplicity and reduced need for stitching [18, 19]. This scanning approach is naturally compatible with confocal line illumination [20]. Stage-scanning instead of beam scanning makes for a system with a minimum number of moving optical components. A line scanning system can be extended by replacing the line sensor with an area sensor. This gives additional freedom for hyper spectral scanning approaches [21, 22].

In this paper we propose a multi-focal multi-line scanning fluorescence microscope for efficient 3D imaging over large, cm^2 sized, scanning areas. Figure 5.1 shows the essentials of the scanner concept. The architecture is based on the use of an area image sensor that is tilted with respect to the optical axis, inspired by [23]. The fluorescent specimen is scanned with a set of illumination lines, that are oriented perpendicular to the plane spanned by the optical axis and the scan direction, are focused at different depths inside the specimen, and that are optically conjugate to rows of pixels of the image sensor. This makes it possible to apply confocal slit apertures digitally, post-acquisition. Another advantage is that the line foci used for scanning are laterally separated, avoiding losses in the collected fluorescence signal by beam splitters in existing approaches [11, 12].

The major challenge in realizing this new scanning concept lies in the area of PSF engineering, in which many approaches targeting the imaging light path, often in the context of single-molecule imaging [24–29], as well as the illumination light path, in particular light sheet based techniques [30, 31], have been proposed. In the current case the single laser beam must be transformed into a set of parallel scan lines projected at different depths inside the sample. Splitting a beam with a diffraction grating into a set of sub-beams with near equal intensity that are focused by the objective lens into a set of equidistant scanning spots in the scan (y) direction is an obvious possibility. In this paper we show how to extend the use of diffractive optical elements to first add defocus such that the set of spots scan the sample at equidistant focal layers, and to second modify the spots to lines orthogonal to the plane spanned by the optical (z) axis and scan (y) axis. Moreover, we introduce a design method for such lines to have uniform intensity along the lines. Secondary challenges in realizing the scanner concept lie in calibration and alignment of the diffractive optical element with the system, in particular with the pixel rows of the image sensor, and with the need to correct for aberrations in the illumination light path. Minimizing the spherical aberration arising from the finite conjugate imaging at different depths is an issue that also must be addressed.

This article has the following structure. In the theory section, the illumination PSF engineering approach to generating the set of equidistant multi-focal scan lines will be described in detail. The experimental methods section starts with a description of our experimental setup. After this, the use of a Spatial Light Modulator (SLM) as diffractive optical element in the setup will be discussed, and, finally, the image metric based aberration correction that is implemented, will be described. In the experimental results section, a quantitative characterization of the realized illumination PSF is presented; images of immunofluorescently stained

cells and tissues and of cells stained with Fluorescence *in situ* Hybridization (FISH) are shown; and a quantitative description of the photo-electron yield of the system, needed to assess the potential optical throughput, is provided. The article is concluded with a discussion of further extension of this scanning platform to higher resolutions and to multi-color imaging.

5.2. Theory

5.2.1. Design for generating a set of defocused points

The first step in engineering the desired illumination pattern is to split the illumination laser beam into a set of sub-beams of near equal power, such that the sub-beams form a set of distinct foci after focusing by the lens. Each focus must be a line and the total set of foci must have an equidistant spacing in the scan (y) direction and in the focus (z) direction. Our design for this is a diffractive structure, often referred to as a 'diffractive optical element'.

Our design approach is based on the method used earlier by one of us [29]. For the sake of completeness, the essentials of the method are repeated here. The diffractive structure is assumed to be placed in a plane conjugate to the pupil plane and is described as a thin surface, locally altering the phase of the incoming light, independent of the angle of incidence. Then the complex amplitude of the incoming beam is modified by the transmission function $T\left(\vec{\rho}\right) = \exp\left(2\pi i W(\vec{\rho})/\lambda\right)$, where $\vec{\rho}$ is the normalized pupil coordinate and $W\left(\vec{\rho}\right)$ is the phase profile added to the incident beam. This phase profile is described by:

$$W\left(\vec{\rho}\right) = f\left(\mathrm{mod}\left[\frac{K\left(\vec{\rho}\right)}{\lambda}\right]\right), \tag{5.1}$$

where λ is the illumination wavelength, $f(t)$ the so-called profile function with $t \in [0,1)$, and $K\left(\vec{\rho}\right)$ the so-called zone function. It appears that the incident beam is split into diffraction orders with integer index m with amplitude:

$$C_m = \int_0^1 dt \exp\left(-2\pi i m t\right) \exp\left(\frac{2\pi i f(t)}{\lambda}\right), \tag{5.2}$$

depending only on the profile function, and phase:

$$W_m\left(\vec{\rho}\right) = mK\left(\vec{\rho}\right), \tag{5.3}$$

depending only on the zone function. For a diffraction grating, the zone function is a linear function of the pupil position, leading to an added aberration that only consists of wavefront tilt. This is sufficient for splitting the incoming beam into a set of diffraction orders (the required set of sub-beams), that will result in a set of foci with an equidistant spacing in the lateral (scan, y) direction. In our case, however, it is desired to have an equidistant spacing of the diffraction orders in the axial (focus, z) direction as well. We therefore choose a zone function that is a sum of tilt and defocus:

$$K\left(\vec{\rho}\right) = \Delta y \frac{\mathrm{NA}\rho_y}{\lambda} + \Delta z \frac{\sqrt{n^2 - \mathrm{NA}^2|\vec{\rho}|^2}}{\lambda}, \tag{5.4}$$

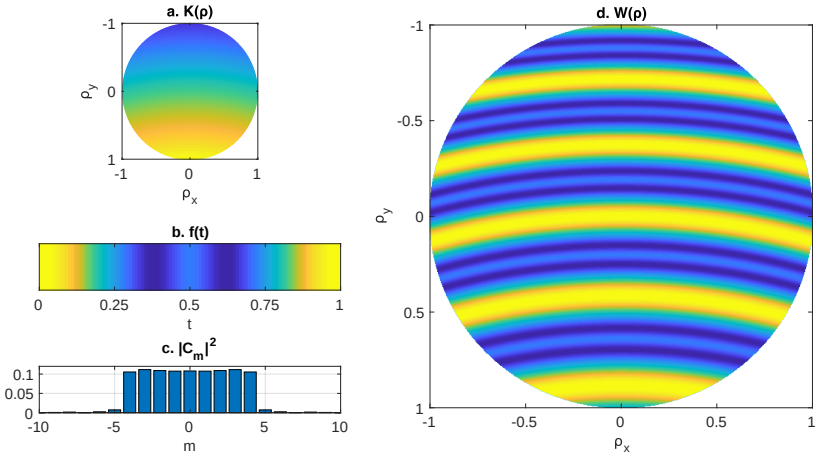


Figure 5.2: Demonstration of grating with tilt and defocus. (a) The zone function with $\Delta y = 10\lambda$ and $\Delta z = 10\lambda$. These numbers are chosen for demonstration purpose, and do not represent a realistic scenario. (b) The profile function optimized for 9 lines. (c) The resulting distribution of power over the diffraction orders. (d) The corresponding aberration function.

with Δy the lateral spacing of the foci, Δz the axial spacing of the foci, n the refractive index of the sample medium, and NA the microscope objective Numerical Aperture.

The next step in the design is to find a profile function $f(t)$ that gives a power distribution over the diffraction orders $|C_m|^2$ that is as uniform as possible. This problem is in general related to the problem of finding a phase function that leads to a desired intensity distribution, and can for example be solved using the Gerchberg-Saxton algorithm [32]. We use a parameterization of the profile function in terms of a finite Fourier series:

$$f(t) = \sum_{n=1}^{N} a_n \cos(2\pi n t), \tag{5.5}$$

which advantageously results in a smooth and band-limited profile function, and is characterized by the N free parameters a_n. The cost function is defined as

$$Z(a_0, \ldots, a_N) = \sum_{m=-M}^{M} \left(|C_m|^2 - \eta_m\right)^2, \tag{5.6}$$

where an integer $M \geq N$ is used, and where η_m represents the desired set of diffraction efficiencies. An even distribution of power over the first $2K+1$ diffraction orders is obtained when $\eta_m = 1/(2K+1)$ for $m \leq K$ and zero otherwise. A generic search algorithm can be used for the minimization problem, such as implemented in the function `fminsearch` in MATLAB. In our experiments we have implemented a solution for $K = 4, N = 5, M = 10$, which results in a profile function that gives 97 % of the power in the 9 diffraction orders, where the power in the individual orders varies by less than 3 %, see Fig. 5.2.

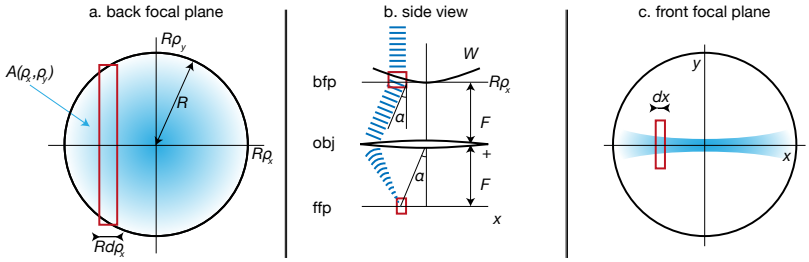


Figure 5.3: Illustration of the proposed geometry for the aspherical cylindrical lens design. (a) In the back focal plane, we have an circular aperture which is illuminated by a beam profile $A(\rho_x, \rho_y)$. We introduce an astigmatic aberration function W such that every strip in the back focal plane there is a corresponding strip in the front focal plane (red boxes in all subfigures). (b) Side view of the optical system, with the back focal plane (bfp), the objective lens (obj) and the front focal plane (ffp). (c) The strips in the front focal plane have a non-constant width dx. By changing the shape of the W, dx can be modified to obtain a uniform distribution of power.

5.2.2. Design for lines with uniform intensity

So far, the diffractive structure is capable of producing a set of point foci, equidistantly spaced in the scan (y) direction as well as in the focus (z) direction. The next step is to modify the set of point foci into a set of line foci, where the direction of the lines is orthogonal to the yz-plane of the original point foci. This can be accomplished by adding astigmatism to the aberration profile, giving a total aberration profile:

$$W\left(\vec{\rho}\right) = W_{yz}\left(\rho_x, \rho_y\right) + W_x\left(\rho_x\right), \tag{5.7}$$

where $W_{yz}\left(\rho_x, \rho_y\right)$ is the aberration profile of the grating structure, according to Eqs. (5.1), (5.4) and (5.5), and where $W_x\left(\rho_x\right)$ represents astigmatism. This aberration function depends only on the normalized pupil coordinate ρ_x as the intended line foci are oriented in the lateral plane orthogonal to the scan (y) direction. The astigmatic aberration can be generated by a refractive optical element such as a cylindrical lens or by the same diffractive optical element that produces the grating structure.

It appears that a simple parabolic astigmatic aberration profile $W_x\left(\rho_x\right) = \frac{1}{2}a{\rho_x}^2$ with a a coefficient, results in a line focus where the peak intensity and line width vary along the line. Here, we add higher order astigmatism in order to overcome these issues, generating focal lines with substantially constant peak intensity and line width along the line. The optical design for these improved astigmatic focal lines starts with a simplified 1D diffraction model for computing the impact of a general aberration profile $W_x\left(\rho_x\right)$ on the focal line shape. Consider an infinitesimal strip in the back focal plane at location ρ_x with width $Rd\rho_x$ and height $2R\sqrt{1-{\rho_x}^2}$, with R the pupil radius as indicated in Fig. 5.3a. This strip generates a plane wave that makes an angle α with the optical axis as indicated in Fig. 5.3b, given by:

$$\sin\alpha = \frac{1}{R}\frac{dW_x}{d\rho_x}. \tag{5.8}$$

This plane wave is focused to a point in the field of view a distance x from the

optical axis given by:

$$x = F\sin\alpha = \frac{1}{\mathrm{NA}}\frac{dW_x}{d\rho_x}, \tag{5.9}$$

with F the focal length of the objective lens, which we assume to be aplanatic and where we have used that $\mathrm{NA} = R/F$. The strip width in the front focal plane as illustrated in Fig. 5.3c is given by:

$$dx = \frac{d(F\sin\alpha)}{d\rho_x}d\rho_x = \frac{1}{\mathrm{NA}}\frac{d^2W_x}{d{\rho_x}^2}d\rho_x. \tag{5.10}$$

Equation (5.9) provides a unique mapping of pupil coordinate ρ_x to field position x if the second order derivative of W_x as a function of ρ_x does not change sign. Now, the intensity in the front focal plane according to the 1D diffraction model is:

$$I_f(x,y) = \left(\frac{1}{\mathrm{NA}}\frac{d^2W_x}{d{\rho_x}^2}\right)^{-1}\cdot\left|U_f(x,y)\right|^2, \tag{5.11}$$

where the first factor is the ratio between the width of the strip in the back focal plane and the width of the strip in the front focal plane. The second factor is the intensity resulting from diffraction in the ρ_y-direction:

$$U_f(x,y) = \frac{\mathrm{NA}}{\lambda}\int_{-\sqrt{1-\rho_x{}^2}}^{\sqrt{1-\rho_x{}^2}} d\rho_y A\left(\rho_x,\rho_y\right)\exp\left(-\frac{2\pi i\rho_y y\mathrm{NA}}{\lambda}\right), \tag{5.12}$$

where the mapping from pupil coordinate ρ_x to field position x is implicitly taken into account, and where $A(\rho_x,\rho_y)$ is the amplitude profile of the beam incident on the diffractive structure. It is mentioned that the 1D diffraction model can also be derived from the standard 2D diffraction integral over the pupil plane if we apply the stationary phase approximation to the integral over the pupil coordinate ρ_x.

Assuming that the incident amplitude profile is flat, $A\left(\rho_x,\rho_y\right) = A$, we find that:

$$U_f\left(x,y\right) = \frac{2A\mathrm{NA}}{\lambda}\sqrt{1-\rho_x{}^2}\,\mathrm{sinc}\left(2\pi\sqrt{1-\rho_x{}^2}\frac{y\mathrm{NA}}{\lambda}\right), \tag{5.13}$$

with $\mathrm{sinc}\,(t) = \sin(t)/t$. The focal line intensity according to Eq. (5.11) is:

$$I_f\left(x,y\right) = \left(\frac{1}{\mathrm{NA}}\frac{d^2W_x}{d{\rho_x}^2}\right)^{-1}\frac{4A^2\mathrm{NA}^2}{\lambda^2}\left(1-\rho_x^2\right)\left[\mathrm{sinc}\left(2\pi\sqrt{1-\rho_x{}^2}\frac{y\mathrm{NA}}{\lambda}\right)\right]^2. \tag{5.14}$$

For the simple parabolic phase profile $W_x\left(\rho_x\right) = \frac{1}{2}a{\rho_x}^2$, corresponding to a cylindrical lens in the paraxial approximation, the pupil coordinate ρ_x and field position x are related by $\rho_x = \mathrm{NA}x/a$, and we find a focal line intensity profile:

$$I_f\left(x,y\right) = \frac{4A^2\mathrm{NA}^3}{a\lambda^2}\left(1-\left[\frac{\mathrm{NA}x}{a}\right]^2\right)\mathrm{sinc}\left(2\pi\sqrt{1-\left[\frac{\mathrm{NA}x}{a}\right]^2}\frac{y\mathrm{NA}}{\lambda}\right)^2. \tag{5.15}$$

The peak intensity decays as $1-(\mathrm{NA}x/a)^2$ along the line, and the line width increases as $1/\sqrt{1-(\mathrm{NA}x/a)^2}$ along the line.

A peak intensity along the line that is independent of x can be achieved if:

$$\frac{d^2 W_x}{d{\rho_x}^2} \propto 1 - {\rho_x}^2. \tag{5.16}$$

This results in an aberration profile $W_x(\rho_x)$ that is of fourth order in ρ_x. Ignoring constant phase offsets, and requiring a symmetric line structure centered at the optical axis (no odd terms in ρ_x), we arrive at:

$$W_x(\rho_x) = \frac{3}{4} w\mathrm{NA}\left[{\rho_x}^2 - \frac{1}{6}{\rho_x}^4\right], \tag{5.17}$$

with l an integration constant. The relation between pupil coordinate ρ_x and field position x according to Eq. (5.9) is:

$$x = \frac{3}{2} l\left[\rho_x - \frac{1}{3}{\rho_x}^3\right]. \tag{5.18}$$

At the pupil rim $\rho_x = 1$ we find that $x = l$, implying that the total line length is $2l$. Eq. (5.18) may be inverted to:

$$\rho_x = 2\cos\left(\frac{\pi}{3} + \frac{1}{3}\arccos\left(\frac{x}{l}\right)\right). \tag{5.19}$$

It is mentioned that, although the peak intensity is constant along the line, the line width still depends somewhat on field position x. An alternative requirement that might be pursued is the requirement that the total power is distributed equally along the line, i.e. $\int dy I_f(x,y)$ is constant. It then follows that:

$$\frac{d^2 W_x}{d{\rho_x}^2} \propto \int dy |U_f(x,y)|^2 \propto \sqrt{1-{\rho_x}^2}, \tag{5.20}$$

which can be solved in a similar way as before, leading to:

$$W_x(\rho_x) = w\frac{2\mathrm{NA}}{\pi}\left[\rho_x \arcsin(\rho_x) - \frac{1}{3}\left(1-{\rho_x}^2\right)^{\frac{3}{2}} + \sqrt{1-{\rho_x}^2}\right] \tag{5.21}$$

In our experiments we have opted for keeping the peak intensity along the line as constant as possible, which is optimal in the limit of a small confocal detection slit.

We have tested the theoretical analysis with our 1D diffraction model by computing the shapes of line foci computed numerically with standard 2D diffraction by a circular aperture. Figure 5.4 shows an evaluation of the obtained results for a design line length $2l = 500\,\lambda/\mathrm{NA}$. Figure 5.4a shows the result for an aberration function that consists of lowest order astigmatism only, as e.g. created by the use of a cylindrical lens. The peak intensity over the line decreases to zero and the line broadens towards the edges. The full width at half maximum of this line is $1.42l$, i.e. just 71 % of the total intended line length $2l$. The practically usable part of the line is even less than this as a drop of 50 % in peak intensity is already quite high. Figure 5.4b shows the result for the aberration function including higher order astigmatism, optimized for uniform peak intensity according to Eq. (5.17). The peak intensity turns out to be uniform over the full line, as predicted by the 1D

diffraction model, except for a small amount of ringing close to the line edges as can be seen in Fig. 5.4c. This ringing originates from diffraction at the aperture edges in the ρ_x-direction, which is not accounted for in the 1D diffraction model. At $x = \pm 220\lambda/\mathrm{NA}$ ringing is no longer present and the line width is within one Airy unit, as shown in Fig. 5.4d. We conclude that about 85 % of the design line length can be used, without a substantial decrease in peak intensity or a substantial increase in line width.

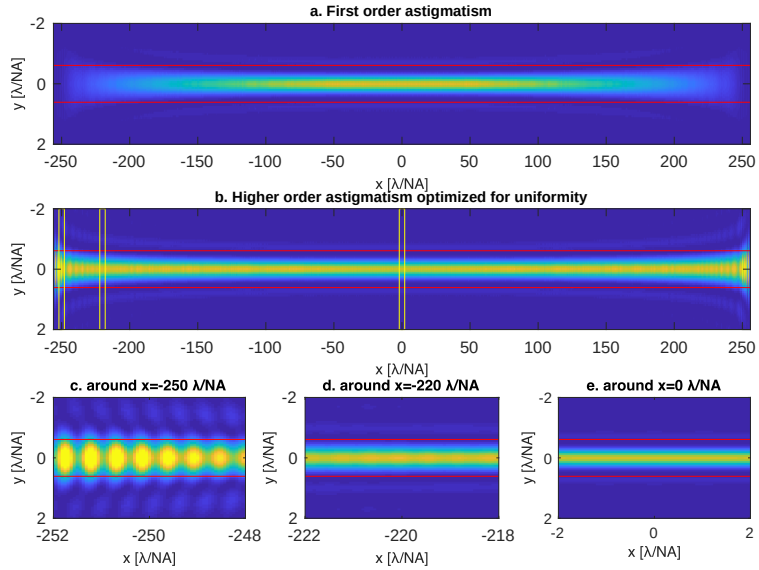


Figure 5.4: Calculated intensities in the focal plane using 2D scalar diffraction. A value $l = 256\lambda/\mathrm{NA}$ was used. The aperture is assumed to be circular. The red lines indicate one airy unit. (a) A line created using a parabolic aberration function. (b) A line created using an aberration function with higher order astigmatism optimized for uniformity conform Equation (5.17). The yellow boxes correspond to the detailed images given in c-e. (c) Close to the edge of the line profile ringing is present. (d) This ringing has disappeared around $x = -220\lambda/\mathrm{NA}$. Here the line width is within one Airy unit. (e) In the center of the line pattern, the line has the smallest line width. The peak is constant over the line.

5.2.3. Image formation theory

Several aspects of the envisioned imaging system have impact on the formal image formation model. First of all, the illumination is with a set of line foci, making the illumination intensity as a function of position in object space:

$$H_{ill}(x,y,z) = \sum_m H_{ex}(y - m\Delta y, z - m\Delta z), \tag{5.22}$$

where $H_{ex}(x,z)$ is the intensity profile of the line foci, Δy is the lateral separation of the line foci, Δz is the axial separation of the line foci, and the sum is over all participating diffraction orders. Second, the image data is captured at a tilted image sensor, which mixes the lateral and axial coordinates. For a fluorescent object $F(x,y,z)$ recorded for a scan coordinate y_s, the intensity recorded at the image

sensor is:

$$I\left(Mx',My',y_s\right)=\int dxdydz\,H_{em}\left(x'-x,y'-y,-z+\frac{\Delta z}{\Delta y}y'\right)F\left(x,y+y_s,z\right)H_{ill}\left(x,y,z\right), \tag{5.23}$$

where $H_{em}\left(x,y,z\right)$ is the emission Point Spread Function (PSF), and M is the lateral magnification. This may be written as a sum over line images:

$$I\left(Mx',My',y_s\right)=\sum_m I_m\left(x',y'-m\Delta y,y_s\right), \tag{5.24}$$

with (substitute $y''=y'-m\Delta y$):

$$\begin{aligned}I_m\left(x',y'',y_s\right)=&\int dxdydz\,H_{em}\left(x'-x,y''-y+y_s,-z+\frac{\Delta z}{\Delta y}y''\right)H_{ex}\left(y-y_s,z\right)\\&\times F\left(x,y+m\Delta y,z+m\Delta z\right),\end{aligned} \tag{5.25}$$

Taking into account that only a small range of y'' values around the central line position at $y''=0$ give rise to substantially non-zero intensities we may approximate this as:

$$\begin{aligned}I_m\left(x',y'',y_s\right)=&\int dxdydz\,H_{em}\left(x'-x,y''-y+y_s,-z\right)H_{ex}\left(y-y_s,z\right)\\&\times F\left(x,y+m\Delta y,z+m\Delta z\right).\end{aligned} \tag{5.26}$$

The procedure of digital spatial filtering following a slit shape, possibly, including the collection of extra signal along the scan with a TDI operation can be incorporated by setting $y_s=y_i-y''$ and integrating over y'', where the integration range is K rows of pixels (we take $K=4$ or $K=1$). Here y_i is the image coordinate in the scan direction. The image coordinate in the field direction is simply found by $x'=x_i$. Then we find for the image data at layer m:

$$\begin{aligned}J_m\left(x_i,y_i\right)=&\int dy''\,I_m\left(x_i,y_i,y_i+y''\right)\\=&\int dxdydz\,H\left(x_i-x,y_i-y,-z\right)F\left(x,y+m\Delta y,z+m\Delta z\right),\end{aligned} \tag{5.27}$$

where the total PSF:

$$H\left(x,y,z\right)=H_{em}\left(x,y,z\right)\left[\int dy''\,H_{ex}\left(-y+y'',-z\right)\right] \tag{5.28}$$

is seen to be the product of the emission PSF with the average of the line excitation PSF averaged over a lateral distance corresponding to K rows of pixels. Close to focus, this integration range is sufficient to capture the entire excitation signal, maximizing signal, and reducing the overall PSF to the emission PSF of the system. This stands in contrast to conventional confocal point scanning microscopes, where the imaging in focus is determined by the excitation PSF. Further away from focus, the integration range is substantially less than the width of the defocused lines. Consequently, the captured signal is reduced and optical sectioning is achieved. The complete sectioning behavior is described by the background function:

$$B\left(z\right)=\int dxdy\,H\left(x,y,z\right), \tag{5.29}$$

which gives the contribution to the background of a layer a distance z away from the imaged plane. For non-sectioning widefield imaging $B(z) = 1$, for sectioning modalities $B(z)$ is peaked around $z = 0$, decaying to zero as $|z| \to \infty$.

The illumination lines are characterized experimentally by imaging thin uniform fluorescent layers, for which $F(x,y,z) \approx \delta(z-z')$, with $\delta(z)$ the delta-function, in a static configuration without any scanning ($y_s = 0$) but with a tunable defocus z'. Then the expected line image is a function of the image coordinate in the scan direction y' and of the defocus z' only:

$$I_m(x',y',z') = \int dxdy H_{em}(x'-x, y'-y, m\Delta z - z')\, H_{ex}(y, z' - m\Delta z), \qquad (5.30)$$

which leads to an image that is the average of the emission PSF in the field (x) direction, and the convolution of the line excitation PSF and the emission PSF in the scan (y) direction. This image function depends is dominated by the broadest of the two functions involved in the convolution, either the line excitation profile or the average of the emission PSF in the field direction. Integrating the obtained images over the lateral coordinate provides a measurement of the background function $B(z)$ needed to assess the degree of confocality.

The illumination lines are further characterized experimentally by imaging thick fluorescent layers for which $F(x,y,z)$ is substantially constant, in a static configuration without any scanning ($y_s = 0$), and with an additional defocus z' applied to the excitation PSF. In that case the expected line image as a function of y' and z' is given by:

$$T_m(y',z') = \int dxdydz H_{em}(x'-x, -y, z'-z)\, H_{ex}(y,z), \qquad (5.31)$$

where a change of integration variable is used for the integration over the axial coordinate. This through-focus line PSF like function is thus the 2D convolution of the through-focus line excitation PSF and the average of the through-focus emission PSF over the field coordinate. The measured through-focus line shape is determined by the broadest of these two functions.

5.3. Experimental methods

5.3.1. Experimental setup

The experimental setup is illustrated in Fig. 5.5. A Nikon 20× NA 0.75 Plan Apochromat VC objective lens with a focal length of $F_{ob} = 10\,\text{mm}$ and a custom Nikon tube lens (effective focal length of $F_{tube} = 222.4 \pm 2.2\,\text{mm}$) are used, giving a lateral magnification of the imaging light path $M = 22.24$. The axial magnification of the imaging path is $M_{ax} = \chi M^2/n = 352$, where a sample refractive index $n = 1.5$ is assumed and $\chi = 1.07$ is a non-paraxial correction factor [33, 34]. The specimen (slide) is scanned using two stages: a PI M-505 low profile translation stage for positioning of the slide in the field direction and a Newport XM1000 ultra precision linear motor stage for the continuous scanning motion of the sample. For focusing two stages are used to axially translate the objective: a PI M-111 compact micro-translation stage for coarse positioning, and a PI P-721.CL0 piezo nanopositioner for fine positioning.

A Hamamatsu Orca Flash 4.0 v2 camera is used with a pixel pitch of 6.5 µm corresponding to $p = 0.29\,\mu m$ in object space, which somewhat undersamples the imaging PSF (Nyquist sampling is at $\lambda/4NA = 0.20\,\mu m$ for $\lambda = 590\,nm$). The gain of the camera is measured to be $0.39\,e^-/ADU$. The camera is tilted over an angle of $\beta = 20°$ with respect to the optical axis. The image sensor size in the direction perpendicular to the line sensors is $2048 \times 6.5\,\mu m = 13.3\,mm$, such that the axial range in image space is $d = 4.4\,mm$ and the axial range in object space is $d/M_{ax} = 12.5\,\mu m$. The Orca Flash camera has a micro-lens array for higher photon detection efficiency, which makes the sensor less efficient for light that is incident at an angle. The measured loss in light collection efficiency is shown in Fig. 5.5b. It appears that tilting the detector 20° reduces the light collection efficiency to 53 % of its nominal value. The camera supports exposure times down to 1 ms. The frame rate, however, is limited by the time required for the rolling shutter to read out the sensor area, leading to a maximum frame rate of 100 fps for a full frame acquisition.

We found the camera to skip some frames when operated at high frame rates, which was solved by reducing the frame rate to 21 fps.

The light source is an Omicron LightHUB-4 equipped with a PhoXX+ 488-100 laser with 100 mW output power and 488 nm wavelength. A polarization maintaining broadband fiber with an achromatic collimator is used to obtain a collimated beam in free space. A beam expander is built using a Thorlabs AC254-50-A, and AC254-200-A achromats for a 4× magnification. A Liquid Crystal on Silicon (LCoS) Spatial Light Modulator (SLM, Boulder Nonlinear Systems XY Phase 512) is used for the required illumination PSF engineering. The SLM has $N^2 = 512 \times 512$ pixels and a square aperture of 7.68 mm, resulting in a pixel pitch of 15 µm. A Thorlabs AHWP10M-600 achromatic half wave plate is used to align the polarization of the

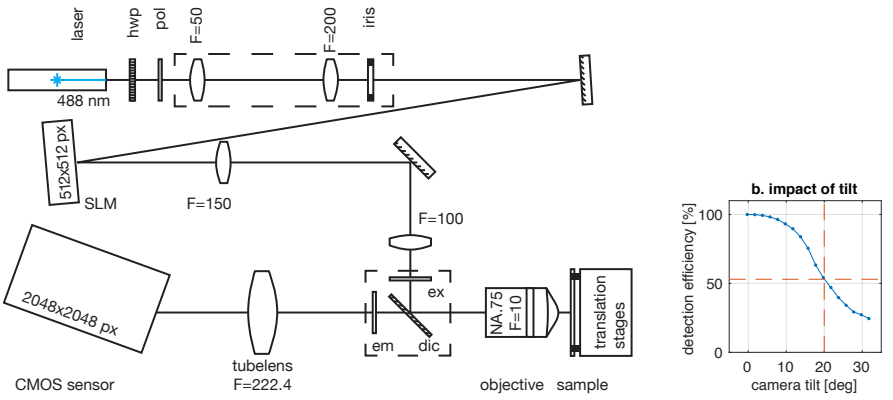


Figure 5.5: Schematic illustration of the multi-focal line confocal scanner prototype. The polarization of the laser beam is aligned to the SLM using a half wave plate (HWP) and filtered using a polarizer (POL). The beam is expanded and projected onto the SLM. The diffracted light is coupled into the objective lens using a filter cube consisting of an excitation filter (EX), emission filter (EM) and a dichroic mirror (DIC) and imaged in the back focal plane of the objective. The fluorescence light is imaged onto the CMOS sensor using the tube lens. Focal lengths F are given in mm. (b). The measured impact on the photon detection efficiency of tilting the camera with respect to the optical axis. For a tilt of 20° the detection efficiency is reduced with 53 % with respect to orthogonal incidence.

beam to the SLM and a Thorlabs LPVISE100-A linear polarizer is used to suppress the orthogonal polarization. A relay path from SLM to the back focal plane of the objective lens is made using a Thorlabs AC254-150-A achromat ($F = 150\,\text{mm}$), and an AC254-100-A achromat ($F = 100\,\text{mm}$ giving a magnification of $M = 0.67$). A Semrock LF405/488/532/635-B-000 quad-band dichroic mirror is used to couple in the illumination light into the objective and direct the emission light towards the camera. A National Instruments NI PXIe-6363 data acquisition card (DAQ) is used for synchronization of the hardware. The DAQ gives a start trigger to the camera for each frame, and it provides a high signal to the laser for digital modulation, such that the laser is only switched on during the camera integration period. The acquisition card is synchronized to the output trigger of the scanning stage.

5.3.2. Fluorescent test samples

Two test slides are used to calibrate and test the scanner setup. First, we use a 1.4 mm thick green plexiglas slide of Chroma that shows strong autofluorescence when excited with blue light. A glass coverslip was attached onto the slide for proper spherical aberration matching. This removes the need for imaging deep into the material, and so prevents scattering and reduces out-of-focus fluorescence. This slide has high fluorescence, is relatively homogeneous and has low bleaching. Second, we used a slide with a spin coated fluorescent layer of Rhodamine on a coverslip. This slide has the advantage of being very thin ($\ll 1\,\mu\text{m}$), but is relatively inhomogeneous and suffers from bleaching.

Two test samples are used to demonstrate the scanner system. The first slide contains a 5 µm thick human prostate tissue micro array (TMA) section labeled using the Kreatech ERBB2 (17q12) / SE 17 FISH probe (product number KBI-10701). This labeling is used for the detection of amplification of the ERBB2 (also known as HER-2/*neu*) gene via Fluorescence *in situ* Hybridization (FISH). As a reference, also the chromosome 17 centromere is labeled using a Satellite Enumeration (SE) probe. The ERBB2 specific FISH probe is direct-labeled with PlatinumBright550. The SE 17 specific FISH probe gene region is direct-labeled with PlatinumBright495. The SE 17 probe is imaged with our setup. The average wavelength of the fluorophore at the detector (taking into account the emission spectrum and the dichroic filter) is 563 nm. The second slide is a stage 3 human rectum cancer sample with immunofluorescence staining. Three different antibodies label Desmin, CD31, and D2-40. In this work, the Desmin protein (labeled with IgG1 M Alexa Fluor 488) is imaged, which is highly expressed in muscle cells. The average wavelength of the fluorophore at the detector (taking into account the emission spectrum and the dichroic filter) is 539 nm.

5.3.3. Data acquisition and processing

Image data is acquired in a 'push broom' scanning fashion [18, 19]. The speed of the scan movement is $v = p\cos\beta/t_l$, where β is the sensor tilt angle, t_l is the line period and p is the pixel pitch in object (sample) space. Every line period the camera captures a full frame which is transferred to the host computer. A number of regions of interest (ROIs) corresponding to the number of illumination lines are selected, the rest of the image data is deleted. The ROIs have an equal size, span the width of the sensor, and have an equal spacing $s_r = Ms_l/\cos\beta$ where s_l is the

tangential illumination line spacing in object space. The selected ROIs consist of either one or four rows of pixels. This functions as a confocal slit detector of width $p\cos\beta$ or $4p\cos\beta$. The data is added in a Time Delayed Integration (TDI) fashion for the four-line acquisition mode, in order to increase signal-to-noise ratio without sacrificing resolution. [35]. The data obtained in this way for the different ROIs correspond to the layers of the simultaneously acquired focal stack. Storage of the image data to the computer memory of layer m is delayed with a time $m = s_r t_l m$ in order to guarantee lateral alignment of the layers of the focal stack. The data is saved to hard disk after completion of a lane scan.

The following steps are taken in post-processing. First, a fine alignment between the layers is done to compensate for residual misalignments by optimizing the product of all layers on a 1024×1024 pixels patch with distinct point sources. Second, the image is resampled with a factor $\cos\beta$ in the scan direction to have an image with square pixels. Third, a constant value is subtracted to correct for detector offset. Finally, flat fielding is applied to compensate for residual intensity variations (down to 50 % at the edge of the region of interest) mainly arising from the Gaussian illumination profile of the back aperture of the objective.

A maximum intensity projection along the z direction is used for a 2D visualization of the 3D dataset. For improved dynamic range, a gamma correction is applied. The gray scale image is mapped to RGB by multiplying the gray level with an RGB triplet based on the calculated filtered fluorophore emission spectrum: (0.55, 1, 0.08) for Platinum495 and (0.28, 1.0, 0,15) for Alexa488.

5.3.4. Spatial Light Modulator

The so-called space-bandwidth product of the illumination light path (product of illuminated FOV and spatial frequency bandwidth of the illumination pattern) is limited by the number of pixels of the SLM N. Suppose the SLM is imaged by the relay lenses to an aperture of width w in the back focal plane of the objective lens. This provides an illumination NA given by $\mathrm{NA}_{ill} = w/2F_0$ with F_0 the focal length of the objective. The maximum spatial frequency that can be generated by the SLM is $N/2w$, which gives rise to field angles η up to $\sin\eta = \lambda N/2w$, resulting in a field of view:

$$\mathrm{FOV}_{ill} = 2F_0 \sin\eta = \frac{N}{2}\frac{\lambda}{\mathrm{NA}_{ill}}. \tag{5.32}$$

In other words, the product between FOV_{ill} and spatial frequency bandwidth $2\mathrm{NA}/\lambda$ is fixed by the number of SLM pixels N. In the resulting trade-off between FOV_{ill} and NA_{ill}, we have chosen for $\mathrm{NA}_{ill} = 0.26$ and $\mathrm{FOV}_{ill} = 488\,\mu\mathrm{m}$.

An illumination line pitch of $s_l = 54.5\,\mu\mathrm{m}$ is used, corresponding to 200 pixel rows on the sensor. The 9 lines span a total distance $8s_l = 436\,\mu\mathrm{m}$, which conveniently fits in the FOV_{ill}. Intentionally some room was left at the edges of the FOV because this area is most sensitive for ghost lines. These lines correspond to spatial frequency content of the aberration profiles higher than the bandwidth defined by the sampling density of the pixel grid of the SLM. These spatial frequencies are folded back into the central field of view.

Using a multi-line imaging approach can potentially lead to optical cross-talk, i.e. fluorescent signal excited by a line beamlet is detected at the pixel rows that are conjugate to the neighboring line beamlet. The sensitivity to cross-talk can be

estimated using a geometrical optics based argument. Suppose a line of interest has its focus at the origin $(y,z) = (0,0)$. Light is collected from within a cone with a half angle $\alpha_{det} = \arcsin(\mathrm{NA}_{det}/n)$, where NA_{det} is the detection NA and n the sample refractive index. The neighboring line is illuminated with a beam focused at $(y,z) = (s_l, -M\tan\beta s_l/M_{ax})$, where β is the sensor tilt, and M and M_{ax} the lateral and axial magnification of the imaging path. This beam has a half angle of $\alpha_{ill} = \arcsin(\mathrm{NA}_{ill}/n)$, where NA_{ill} is the illumination NA. The z position at which a fluorophore excited by the neighboring illumination beam first enters the cone of the detection NA is given by:

$$z_c = s_l \frac{1 - \tan\alpha_{ill}\tan\beta}{\tan\alpha_{ill} + \tan\alpha_{det}}. \tag{5.33}$$

Filling in the relevant parameters results in $z_c = 72\,\mu\mathrm{m}$. Light from such a distance away from focus will in practice be completely suppressed by the confocal detection.

An SLM has a limited diffraction efficiency, which leaves a part of the incoming beam unmodulated, resulting in a focused 'zero-order' spot in the sample. The diffraction efficiency is affected by several factors [36]. First, the SLM only works for one polarization direction. A half wave plate and a polarizer were included in the beam expander for polarization alignment and suppression of residual orthogonally polarized light. Additionally, the top of the glass surface will contribute to a plane wave reflection, even though it is coated with an anti-reflection layer. Finally, the limited fill factor of the SLM pixels leave parts of the aperture unmodulated. The line illumination pattern is aligned such that the zero-order spot is centered between two lines to enable spatial filtering at the image sensor, and will therefore have no impact on the final 3D image.

The SLM response curve is described by the phase change $\phi(\epsilon)$ as a function of the digital input gray level $\epsilon \in [0,1)$. A lookup table is required to find the gray level given the desired phase shift. A single lookup table is used across the field of the SLM, implying that non-uniformity of the response function is not taken into account. Coarse pixel-to-pixel variations result in low order aberrations that are corrected later on, fine pixel-to-pixel variations result in a small scatter background that has little impact on the final image.

An image based method is developed to obtain the response curve of the SLM. A binary grating with a period much larger than a single SLM pixel and gray levels ϵ_1 and ϵ_2 is created, for a range of values (ϵ_1, ϵ_2). The power in the diffracted orders is obtained by imaging a diffraction pattern that is projected onto the uniformly fluorescent calibration slide. The space (ϵ_1, ϵ_2) is sampled randomly instead of linearly to prevent biases arising from bleaching during data acquisition. Then, the ratio between the power in the first and second diffraction order is measured and fitted with the theoretically expected value:

$$R(\epsilon_1,\epsilon_2) = \frac{\sin\left(\frac{1}{2}\left[\phi(\epsilon_1) - \phi(\epsilon_2)\right]\right)}{\cos\left(\frac{1}{2}\left[\phi(\epsilon_1) - \phi(\epsilon_2)\right]\right) + C}, \tag{5.34}$$

where an extra degree of freedom C is introduced to account for the above mentioned zero order spot and possible background and detector offset. The response

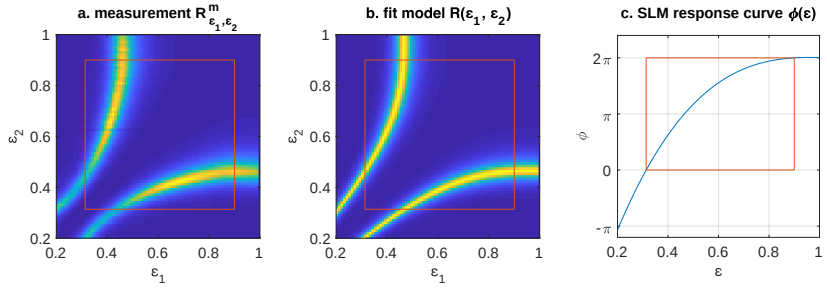


Figure 5.6: Results of the image based calibration of the SLM response curve. (a) Measured ratio between the first and zeroth diffraction order as a function of the gray levels ϵ_1 and ϵ_2 of the binary grating structure. (b) Modeled diffraction order ratio for the found fit parameters. (c) The induced phase change as a function of the digital gray level. The red boxes indicate the gray values that are used to create the lookup table and correspond to a 2π phase change.

function $\phi(\epsilon)$ is parameterized by a polynomial series of order L, and MATLABs `fminsearch` algorithm is used to fit the polynome coefficients. Figure 5.6a shows the measured values, Fig. 5.6b shows the fit model for the obtained fit parameters, and Fig. 5.6c shows the corresponding SLM response curve $\phi(\epsilon)$. We found $L = 3$ to provide enough degrees of freedom to fit the data. The total phase modulation of the SLM exceeds 2π for the used wavelength, which leaves the freedom to choose a subset of ϵ values to create the lookup table. We choose the range between $\epsilon = 0.31$ and $\epsilon = 0.90$, as indicated with the red boxes.

5.3.5. Alignment of Spatial Light Modulator

A careful alignment of the illumination PSF, consisting of the set of multi-focal scan lines, with the pixel grid of the image sensor is required for implementing digital slit pinholes at the imaged line positions in the sensor plane. The reference coordinate system to align the scan lines is given by the optical axis of the objective lens (z-axis) and the two principal directions of the pixel grid of the sensor, which correspond to the field axis (x-axis) and to the scan axis (y-axis). The phase pattern on the SLM needs to be aligned to this coordinate system carefully. A final alignment step is implemented, after mechanical alignment of the SLM in the illumination light path, by means of an affine coordinate transformation between the coordinates of the SLM pixel grid and the pupil coordinates in the back focal plane of the objective lens. In this way three types of misalignment can be corrected. First, the in-plane rotation can be corrected, so that the scan lines are imaged on the sensor parallel to the field axis defined by the pixel grid. Second, the SLM is slightly tilted with respect to the optical axis to separate the incoming and outgoing beam by 11°, see Fig. 5.5. Projection of the tilted sensor plane to the plane perpendicular to the optical axis, distorts the circular beam profile to an elliptical one. This effect may be compensated by an anisotropic stretch in the affine correction transformation. Third, a slight misalignment of the origin of the SLM can be compensated for by a shift of the grating pattern. A small misalignment in the axial position of the SLM can be corrected by adding a small amount of defocus to the induced aberration.

The affine transformation between the coordinates in the SLM plane $\vec{\sigma}$ and the

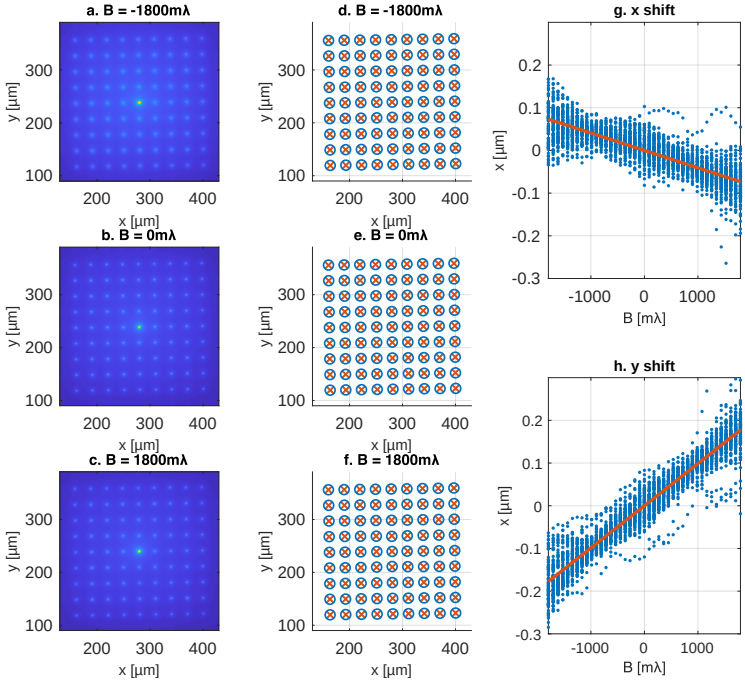


Figure 5.7: Illustration of the alignment calibration using a spot grid. (a-c) image data for three different B values. (d-f) The found spot positions (blue circles) and the fitted model (red crosses). (g-h) The change in the x and y spot coordinate with respect to their average value as a function of B (blue) and the fitted change in position (red).

pupil coordinates $\vec{\rho}$ is given by:

$$\vec{\rho} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} M_x & 0 \\ 0 & M_y \end{bmatrix} (\vec{\sigma} - \vec{\sigma}_0), \tag{5.35}$$

where the first matrix describes the rotation over an angle θ around the z-axis, the second matrix represents the (anisotropic) scaling due to the tilt of the SLM in the light path and due to the overall magnification from the SLM plane to the back focal plane of the objective, and $\vec{\sigma}_0$ is the position offset with respect to the optical axis.

The following aberration profile is fed to the SLM:

$$W = g\left(\mathrm{mod}\left[\frac{1}{\lambda} A \frac{\sigma_x}{R_s}\right]\right) + g\left(\mathrm{mod}\left[\frac{1}{\lambda} A \frac{\sigma_y}{R_s}\right]\right) + \frac{1}{2} B \frac{|\vec{\sigma}|^2}{R_S^2}, \tag{5.36}$$

where g is a pattern function, λ is the wavelength, R_s is the radius of the illuminated area on the SLM, and A and B are freely configurable parameters in units of λ. The first two terms will diffract the light along the horizontal and vertical directions, respectively. The pattern function g is chosen such that light is uniformly distributed over the first 9 diffraction orders (-4$^{\text{th}}$ to +4$^{\text{th}}$ order). The third term will

introduce a defocus, but for nonzero $\vec{\sigma}_0$ it also introduces a lateral shift [28]. The measurement is repeated using a range of 41 values of B spaced with $\Delta B = 90\,\mathrm{m}\lambda$. This aberration profile results in a grid of 9×9 spots. The position of spot (k,l,m), with $k=-4,\ldots,4$, $l=-4,\ldots,4$, and $m=-20,\ldots,20$ in the front focal plane is then given by:

$$\vec{x}_{klm} = \frac{F_0}{R_s}\begin{bmatrix}\cos\theta & -\sin\theta \\ \sin\theta & \cos\theta\end{bmatrix}\begin{bmatrix}1/M_x & 0 \\ 0 & 1/M_y\end{bmatrix}\left(A\begin{bmatrix}k \\ l\end{bmatrix} + m\Delta B\frac{\vec{\sigma}_0}{R_s}\right) + \vec{x}_0, \qquad (5.37)$$

where $\vec{x}_0$ is an overall offset with respect to the optical axis.

The image data and the spot positions of three out of the 41 measurements are shown in Figs. 5.7a and 5.7b, respectively. The change of the spot positions as a function of B with respect to their average position is shown in Fig. 5.7c. A linear least squares fit is used to fit the model of Eq. (5.37) to the measured spot positions. The fit yields the values $\theta = 0.8°, M_x = 1.014, M_y = 1.015, \vec{\sigma}_0/R_s = (-0.03, 0, 08)$. These values are small, as is expected of a mechanically well aligned system, but in need of correction nevertheless.

5.3.6. Correction of aberrations in the illumination light path

The designed illumination PSF, comprising the set of multi-focal scan lines, will be distorted by unintended aberrations in the illumination light path. These can arise from a (coarse) non-uniformity of the SLM pixel response function, or by the components and/or misalignment of the beam expander and the relay path between the SLM and the objective back focal plane. Fortunately, the SLM that is used to introduce intentional aberrations, can also be used to remove the unintentional aberrations. To this end, an image based method is used, without the

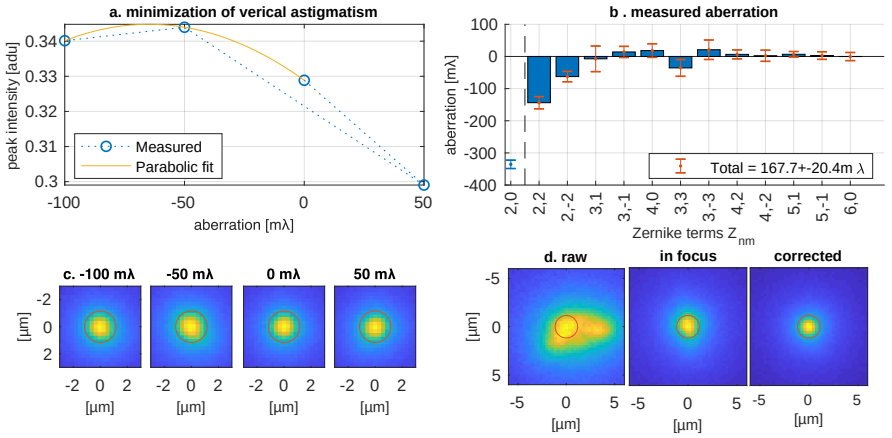


Figure 5.8: (a) Example showing the measurements performed to minimize vertical astigmatism. Four measurements where required to coarsely find the optimum (blue circles). From these measurements a fine estimation of the optimum was calculated using parabolic interpolation (yellow). (b) Graph showing the contributions of the Zernike polynomials to the total aberration. The error bars indicate two standard deviations. (c) The raw image data underlying the vertical astigmatism measurement. (d) The depth averaged illumination spot without any corrections, after minimizing the contribution of Z_{20}, and after aberration correction. The red circles indicate one Airy unit.

need for any additional optical components in the setup [37]. The optimization metric that is used is the peak intensity of a single imaged spot (Strehl ratio optimization). A focused spot is imaged using the thick uniformly fluorescent calibration slide, which is aligned such that the object plane lies about 10 µm deep in the material. In this way a depth averaged illumination PSF is imaged, but this is not problematic for the intended aberration retrieval. The actual peak intensity is determined from a two dimensional parabolic fit of the pixel values around the peak. The to-be corrected aberration profile is parameterized as a sum of orthogonal Zernike polynomials Z_{nm} with radial index n and azimuthal index m, where all modes with $n+|m| \leq 6$ are taken into account, except for piston, tip and tilt. This takes primary and secondary astigmatism, coma, and spherical aberration into account, as well as primary trefoil. Figure 5.8 illustrates the search algorithm used to find the optimal aberration. For each Zernike mode a set of typically three to five spots are recorded with an aberration coefficient that is sequentially changed by an amount $\pm\Delta A$, where the sign is set such that a local optimum in the fitted peak intensity is within the range of recorded aberration settings. A value $\Delta A = 50\,\mathrm{m}\lambda$ is found to properly sample the optimization metric, while only a small number of measurements is required to find the optimum. The optimum aberration value is found by parabolic fitting of the set of fitted peak intensities, where only the maximum value and its two neighbors are included. Figure 5.8a illustrates this search method for primary vertical astigmatism. In this case, four measurements where needed to find the optimum. The corresponding spots are displayed in Fig. 5.8c. This procedure is repeated for all Zernike modes. The result is shown in Fig. 5.8b. Confidence intervals were obtained by repeating the whole procedure ten times on different parts of the sample. Apart from the contribution of Z_{20} (defocus), the total root mean squared (rms) aberration was found to be $168 \pm 20\,\mathrm{m}\lambda$, composed mainly of astigmatism, and significantly higher than the diffraction limit of $72\,\mathrm{m}\lambda$. Figure 5.8d show the illumination spot without any correction, after minimizing the contribution of Z_{20}, and after aberration correction. Clearly, correction of unintended aberrations is a necessity in our setup.

5.4. Experimental results

5.4.1. Characterization of the illumination pattern

The line profile of the illumination PSF is tested experimentally using the thin uniformly fluorescent test slide. The thin fluorescent layer is aligned such that the central illumination line is in focus. The sample translation stage is then used to slowly move the slide during image acquisition in order to create a small 'motion blur' for smoothing out sample irregularities and for spreading the impact of photo-bleaching. The measurement is repeated ten times and the resulting images are averaged for further reduction of noise and sample irregularities. According to the image formation model the expected line shapes are the convolution of the line excitation PSF at focus $m\Delta z$ and the average of the emission PSF in the field direction, at focus $-m\Delta z$. The difference between the imaging NA (equal to 0.75) and the line illumination NA (equal to 0.26) implies that near focus the width of the line excitation PSF is much larger than the width of the emission PSF, whereas further away from focus this is the opposite. The reason is that the relatively small depth of focus of the emission PSF results in a steep increase in the lateral width

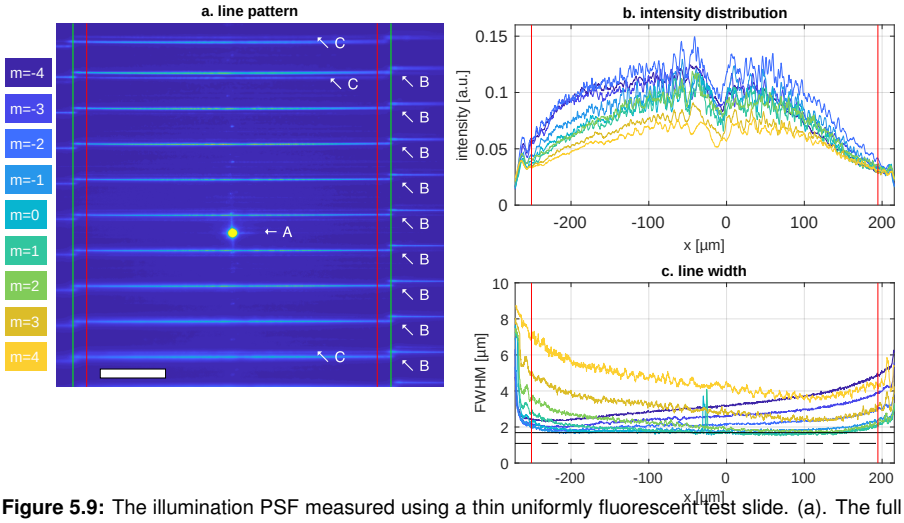


Figure 5.9: The illumination PSF measured using a thin uniformly fluorescent test slide. (a). The full illumination PSF. The scale bar represents 100 µm. The focal lines are labeled with their diffraction order number m and with a color ranging from blue for $m = -4$ to yellow for $m = 4$. The red lines indicate the width of the regions of interest (ROIs) that are used for imaging. The design line length $2l = 488\,\mu\text{m}$ is indicated by green lines. Further indicated are the SLM's zeroth order spot (A), the first alias to the right of the central pattern (B), and higher order focal lines ($|m| > 4$) that are folded back into the central domain due to undersampling (C). These are minor artifacts that are blocked by the spatial slit filter in the digital domain. (b) The intensity distribution over the lines. The higher order astigmatism successfully prevents a zero rim intensity. A decrease of intensity is observed around $x = 0\,\mu\text{m}$ which is caused by an uneven illumination of the SLM. (c) The full width at half maximum (FWHM) of the focal lines. The center lines are in focus and approach the theoretical limit (full black line), a bit broader than the height of a ROI for four line TDI (dashed black line). The line colors in (b) and (c) correspond to the diffraction orders as indicated at the left side of the figure, the width of the x-axis is scaled to the design length $2l$, and the red lines indicate the ROI width.

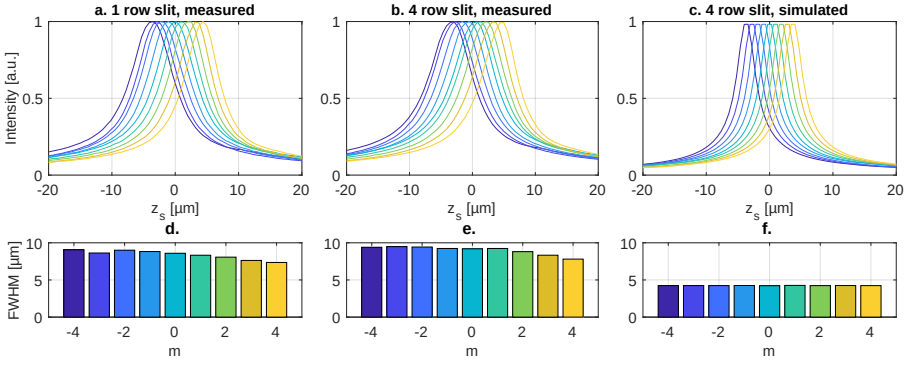


Figure 5.10: Confocal detection of a uniform fluorescent plane. The top row (a,b,c) shows the signal as a function of the sample position in air, for the 9 illumination lines. The bottom row (d,e,f) shows the corresponding axial FWHM. The focal lines are labeled with their diffraction order number m and with a color ranging from blue for $m = -4$ to yellow for $m = 4$. (a, b) The measured signal, while a single row on the sensor is used as a confocal detection. (c, d), The measured response for a four-row confocal detection. (e, f) The simulated response for a four-row confocal detection.

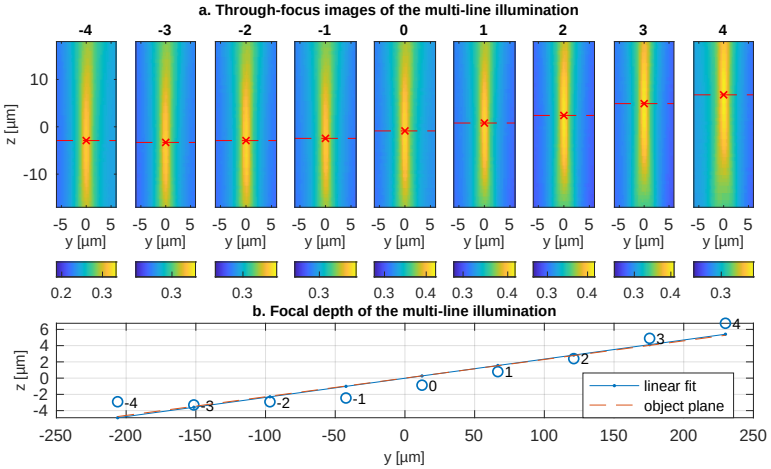


Figure 5.11: Assessment of illumination PSF in a thick fluorescent material. (a) Through focus stack averaged along the x direction. The red crosses mark the focal points. The colorbars below the images indicate the intensity value mapping that is adjusted to cover the full dynamic range in each image. (b) The focal points plotted in the (y,z) plane (blue circles) including a linear fit through the focal points (blue) which overlaps the object plane (red dashed curve).

with defocus.

Figure 5.9a shows the full FOV of the image sensor with the illumination PSF, showing the successfully creation of 9 equally spaced focal lines (labeled with their diffraction order number m). Figure 5.9b shows the intensities of the focal lines along the field direction. The higher order astigmatism based design ensures a substantial illumination intensity over the whole line length. The remaining decrease of power towards the edges is caused by the Gaussian beam profile at the objective pupil. In practice, a decrease of intensity at the edges of the lines to about 30 % is observed. The lines show irregularity on a small length scale resulting from laser speckle. Further, a significant decrease in intensity is visible around $x = 0\,\mu\text{m}$ for all lines, resulting from an uneven illumination of the SLM. Figure 5.9c displays the FWHM of the lines. The center lines are in focus and approach the theoretical limit of FWHM $= 0.89\lambda/\text{NA}$. The FWHM of the outer lines ($m = -4,-3,3,4$) are substantially larger, as they are recorded out of focus, in agreement with the image formation model. A slight tilt of the sample in the field direction is apparent from the fact that the lines with $m > 0$ have a decreasing FWHM for larger x and vice versa for the lines with $m < 0$.

The sectioning capability of the system is tested by measuring the background function as defined in Eq. (5.29). This is done by repeating the previous measurement for a series of axial sample stage positions. The 10 times averaging is omitted for this measurement to limit the illumination photon dose and so avoid bleaching of the sample. The measured images are integrated over the digital slit in the y (scan) direction (up to 4 pixel rows) and over 50 pixels about halfway the line center and the line edge (at around $x = w/4$) in the x (lateral) direction, for noise suppression. As a result, we obtain a background signal as a function of the stage position z for all 9 lines, see Fig. 5.10. For the first result, in Fig. 5.10a, a single row on the

sensor is used as confocal detector. The background function has an average full width at half maximum $\mathrm{FWHM}_z = 8.4\,\mu\mathrm{m}$. In Fig. 5.10b, the result is given the case of a four row confocal detector, which slightly increases the FWHM_z to 9.0 µm on average. The theoretical values for the response to a uniformly fluorescent thin slide are given in Fig. 5.10e. Qualitatively similar curves are found, but less wide than the experimental curves. The FWHM_z is found to be 4.2 µm. The background measured at the bottom scan plane from a source located at the top scan plane is about 20 %, where from the simulation a reduction to about 10 % would have been expected. This result is based on a numerical simulation that computes the illumination PSF H_{ill}, given a defocus following from a stage position that is z_s out-of-focus, according to Eq. (5.22). Then, the intensity at the sensor I is calculated by convolution of H_{ill} with the imaging PSF corrected for a stage defocus of z_s, see Eq. (5.23). The simulated intensity I is saved to disk as an image file and analyzed in the same way as the experimental results. The discrepancy between the width of the measured peak in the background function and the theoretical expectation may be attributed to misalignment, the finite thickness of the layer, and residual aberrations, mainly spherical aberration. The use of scalar diffraction for computing the PSFs instead of more realistic vector PSF models could also lead to a predicted background function peak width that is too optimistic.

The illumination PSF is further tested for uniform spacing and for equidistant foci. To this end, the thick uniformly fluorescent test slide is aligned such that the focal point of the objective is about 10 µm deep in the material. A through-focus stack is created by repeatedly adding a defocus aberration W to the SLM and capturing an image. According to the image formation model the expected line shape is the 2D convolution of the through-focus line excitation PSF and the average of the through-focus emission PSF. Because of the much smaller illumination NA this implies that the measured through-focus line profiles are largely determined by the through-focus line excitation profiles, and hardly by the through-focus emission PSF. Consequently, by using an aberration $W = z\sqrt{n^2 - \rho^2 \mathrm{NA}^2}$, the illumination PSF is sampled at a depth z.

Figure 5.11a shows the measured through focus stacks for an illumination PSF of the 9 beamlets, averaged across the x (field) direction. To find the focus, for every beamlet the position of the maximum intensity is determined. Figure 5.11b displays these positions in the (y, z) plane. The positions have an equal spacing in the y direction of 54.5 µm or 200 pixels on the sensor with a deviation less than 0.14 µm (0.5 pixel). In the z direction the sub beams do have the desired tilted focus. The graph includes a linear fit through the focal points, which overlaps with the tilted object plane with a deviation less than 0.2 µm. However, on top of the linear focus shift, we observe an apparent field curvature leading to a defocus of up to 1.8 µm at the edges. This leads to a slight axial misalignment of the illumination PSF compared to the reference axial imaging depth defined by the imaging PSF. The imaging depth will, however, only be slightly biased as the NA of the illumination is lower than the NA of the imaging path, and therefore the latter will dominate the imaging depth.

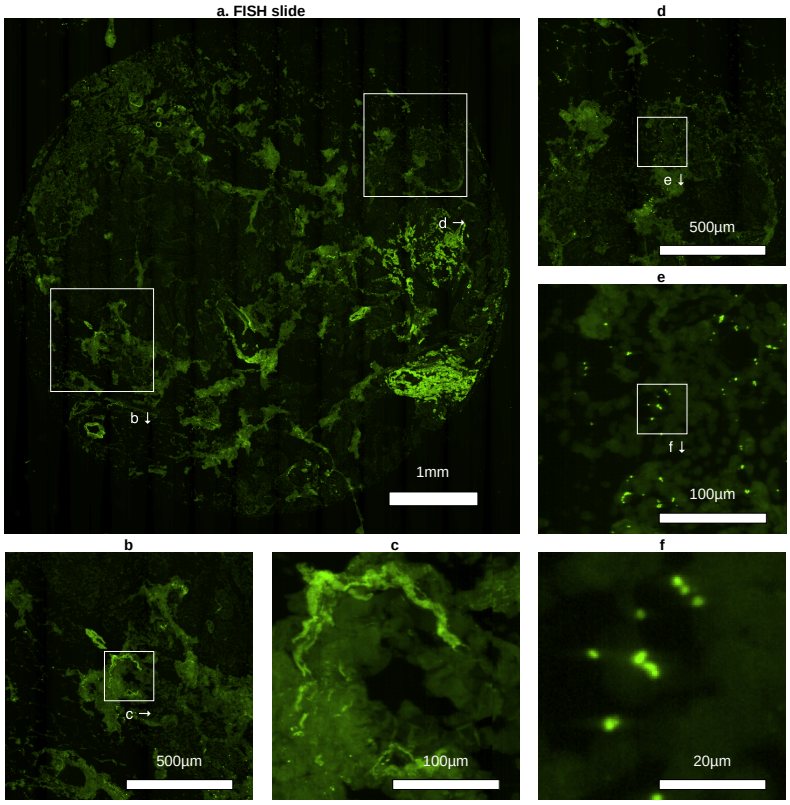


Figure 5.12: A 400 Mpixel maximum intensity projection of a 9 layer multi-focal scan of a tissue micro array (TMA) section labeled using FISH. Every figure zooms in a factor of five, such that the images are shown over a factor of 125 different length scales. The scan is made using an exposure time of 1 ms and four TDI lines. A gamma of 0.7 is used for an enhanced dynamic range. This image is also shown over 100 length scales in Visualization 1*. (a) Whole slide scan. The white boxes indicate two regions that displayed in more detail. (b,d) 5× enlarged. (c,e) 25× enlarged. (f) 125× enlarged.
* doi.org/10.6084/m9.figshare.9912824

5.4.2. Scan results

We present three scan result as proof of concept. First, a 9 layer multi-focal whole slide scan is created of the FISH labeled TMA tissue section. Figure 5.12a shows a maximum intensity projection of the whole slide scan. The image is composed of thirteen scan lanes of 20,000 lines with 1536 pixels width. After resampling this results in a 3600 Mpixel volumetric image ($19,968 \times 18,794 \times 9$ pixels, size $5,790 \times 5,450 \times 11\,\mu m$). For this acquisition, an exposure time of 1 ms and four TDI lines are used. Figures 5.12b to 5.12f provide closer details of the scan and show that zooming in to the cellular level for inspecting tissue structure in Fig. 5.12c and resolving individual FISH spots in Fig. 5.12f is possible.

The second scan result displayed in Fig. 5.13 shows the multi-focal imaging capability of the system. The same FISH labeled tissue is scanned using a single line detector (i.e. no TDI) for better confocality. Here, an exposure time of 10 ms is used. Figure 5.13a shows a frame of a 3D visualization of the z-stack, created

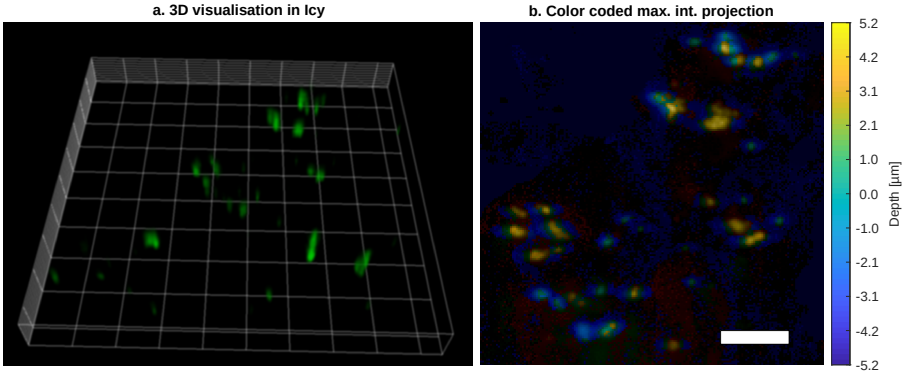


Figure 5.13: 3D visualizations of a 9 layer multi-focal scan of a tissue micro array (TMA) section labeled using FISH. The scan is made using an exposure time of 10 ms and a single line detector. (a) A frame from a 3D visualization using the Icy software [38] (see Visualization 2*). The white line grid has a 10 µm spacing. (b) A maximum intensity projection, where pixels have a color assigned according to the layer of maximum intensity, ranging from –5.2 µm (blue) to 5.2 µm (yellow). The FISH probes can be observed to have different positions in z. The scale bar represents 10 µm.
* doi.org/10.6084/m9.figshare.9912833

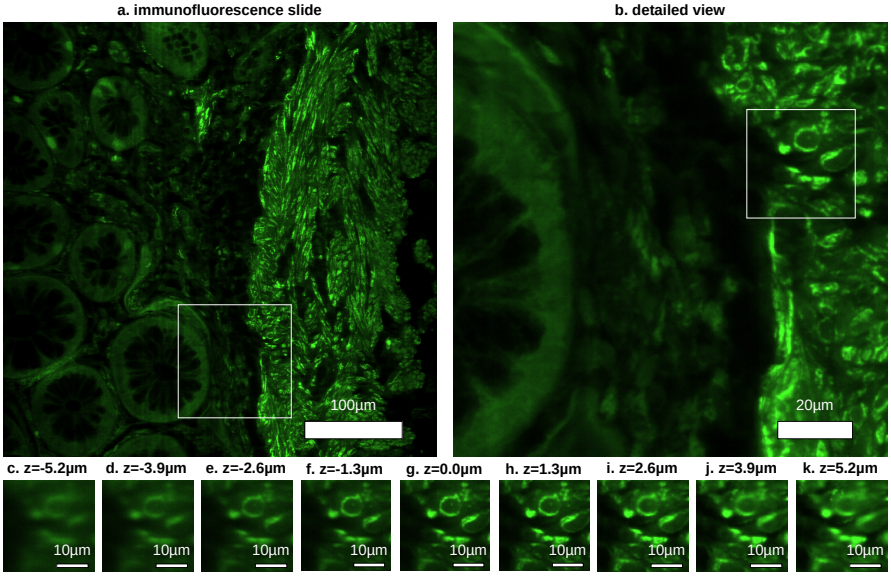


Figure 5.14: A single lane, 9 layer multi-focal scan of an immunofluorescently labeled human rectum tissue section. The scan is made using an exposure time of 10 ms and a single line digital spatial filter. (a) Central layer of the scan, showing the full width of a scan lane. The white box indicates the area that is showed in detail. (b) Detailed view. The white box marks the area displayed in (c-k), which provides a through focus stack of all layers, demonstrating the optical sectioning capability. Although this is a thick tissue section, a high contrast is observed in the central layers.

using the image analysis software Icy [38]. FISH spots can be observed to have different z-positions. This is further illustrated by the color coded maximum intensity projection shown in Fig. 5.13b. In this image, the luminescence of each pixel corresponds to the maximum intensity throughout the layers, while the color corresponds to the layer of maximum intensity. Most spots color orange, indicating a depth of about 3 µm. Some are observed to lay deeper in the sample, ranging from –1 µm (green) to –3 µm.

The third result is a scan of the stage 3 human rectum cancer sample with immunofluorescence staining (10 ms exposure time and using a single pixel row as digital slit). Figure 5.14 shows the central layer of a single lane scan. Figure 5.14b gives a detailed view which provides a good benchmark for the resolution and contrast delivered by the system.

For all scanned samples the reduction of out-of-focus background light by the digital slit pinhole is moderate, in agreement with the measured background function with FWHM_z of around 9 µm, and the thickness of the samples of around 5 µm.

5.4.3. Throughput and resolution

The throughput of the current test setup is $T = N_z N_x f$, with $N_z = 9$ the number of focus layers, $N_x = 1536$ the line length in pixels, and $f = 21$ lines/s the line rate, leading to a modest $T = 0.29$ Mpixels/s. This number is primarily limited by the practical constraints of the image sensor, as the actual exposure time in TDI-mode was just 1 ms, implying a potential gain in throughput with a factor of about 50. A second practical limitation is in the illumination light path efficiency. The main losses were measured to be the coupling into the fiber (35 % loss), the polarizer (35 % loss), the iris that cuts the beam to about one sigma (39 % loss) and the SLM (48 % loss), accumulating to a total illumination light path efficiency of 14 %. These losses are not fundamental and can be improved upon, for example by better alignment of the half wave plate to the polarizer; using free space optics instead of a fiber; a beam shaping optical element to reshape the Gaussian beam into a top hat beam [39]; and by using a dedicated diffractive structure etched in quartz or embossed in a polymer instead of the SLM. A third limitation is the reduced detection efficiency of the camera for light incident at an angle due to the micro lens array (47 % loss). Overcoming these practical hurdles can improve throughput with a factor ranging between 200 and 600 without affecting the SNR. The exposure time could be reduced by an increase in laser power to a decent but not excessive 500 mW. This would further enhance the throughput with a factor of 5. As a result, the proposed optical architecture is estimated to support image acquisitions with a throughput on the order of several hundred Mpixel/s. In the current setup, that would correspond to a scanning speed on the order of 10 mm/s. Note that mechanics do not pose practical constrains on the throughput as the only movement in the system is the linear translation of the sample.

The fundamental limitation to throughput of the proposed scanning architecture is the minimum SNR that is required [40]. The SNR is dominated by shot noise, and therefore the achievable throughput is governed by the minimum amount of photons that must be captured. The maximum line rate f_{max} is proportional to $\eta K P / N_{\min}$, with η the overall photon efficiency of the illumination and imaging light path and detector, K the number of TDI-lines, P the laser power, and $N_{\min}$

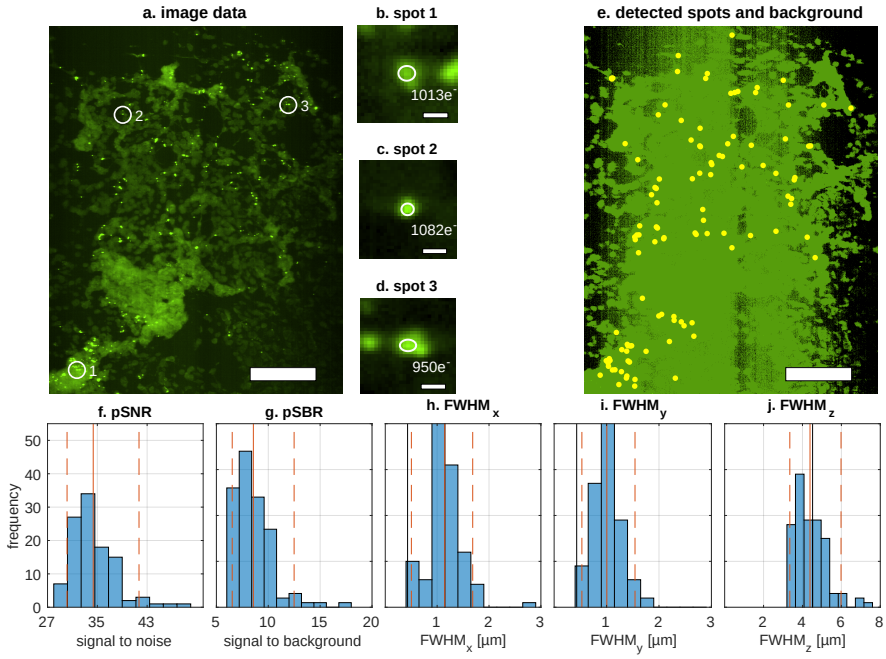


Figure 5.15: Determination of the peak signal to noise ratio (pSNR), the peak signal to background ratio (pSBR), and the lateral and axial resolution expressed using the full width at half maximum (FWHM). (a) Maximum intensity projection of the image data of a scan of the FISH slide made using 1 ms exposure time and a four-line TDI acquisition. The scale bar indicates 100 µm. A gamma correction of 0.7 is applied. Three locations are highlighted. (b-d) A zoomed in view to the highlighted areas. The layer with best focus is displayed. The spots location are given including their peak intensity in photo electrons (e^-). The scale bar indicates 2 µm. The yellow ellipses have a horizontal and vertical diameter equal to the FWHM along the and direction respectively. (e) The image data was classified as background (black) and non-background values (green) or part of a spot. In total, 109 spots where segmented and their location is indicated with a yellow dot. (f-j) Histograms of the pSNR, pSBR and the FWHM along the x, and y direction. The red line indicates the mean and the red dashed lines the 5^{th} and 95^{th} percentile respectively. In (h-j) a black line is included indicating the FWHM of the imaging PSF.

the minimum photon count for a reference fluorescent object.

The relation between the throughput and SNR that is currently achieved is further assessed on the basis of the scan result of the whole slide scan of the FISH slide, see Fig. 5.12. This 5×5 mm tissue section is scanned in 12 lanes of about 17,000 lines, which gives a total camera time (exposure time times number of lines) of less than 3.5 min. We have measured the peak signal to noise ratio (pSNR) and the peak signal to background ratio (pSBR) in a region of 445 µm by 558 µm with a dense distribution of FISH dots. A maximum intensity projection of this region is shown in Fig. 5.15a. A zoomed-in view of three FISH spots are shown in Figs. 5.15b to 5.15d, where we selected the layer with best focus. A classification is performed in which each pixel is classified as being part of the background, part of a spot or neither of both. The classification is done on the 2D maximum intensity projection. The 20^{th} percentile lowest values displayed in Fig. 5.15e are classified as background. An average background level of $139 \pm 7\,e^-$ is found. The pixels with the 0.5^{th} percentile highest are classified as part of a spot or cluster of spots. These pixels are segmented based on 4-connectivity. Small segments are rejected by requiring the shape to span at least three columns and three rows. The peak intensity in every segment was determined by first selecting the layer with best focus and then by interpolating the peak intensity using a two dimensional parabolic fit of the values surrounding the pixel with the maximum intensity to the the model $I(x,y) = w_x(x-x_c)^2 + w_y(y-y_c)^2 + I_p$, where I is the image intensity, w_x, w_y, x_c and y_c are fit parameters and I_P provides the peak intensity. In total, 109 segments are found with peak values ranging from 908 to 1739 e^-. For every segment the pSBR is calculated by dividing the peak intensity by the average background level. Figure 5.15f shows a histogram of the obtained values. An average pSBR of 8.4 (6.5 - 12.0) is found. The pSNR for every segment is approximated as the square root of the peak intensity, as the noise is dominated by shot noise. Figure 5.15g shows a histogram of the obtained pSNR values. An average of 34 (27 - 38) is found. The pSNR and pSBR values that are found are sufficient for state-of-the-art spot segmentation algorithms, which require an SNR level around 5 [41]. This implies that in principle the scan speed could have been 7 times faster.

The lateral and axial resolution of the system are assessed based on the same dataset. Close to the maximum, the spot intensity is well approximated by a two dimensional parabola. Consequently, the lateral full width at half maximum (FWHM) is estimated from the fit parameters w_x, w_y according to $\text{FWHM}_{(x,y)} = 2\sqrt{I_p/2\,w_{(x,y)}}$. Histograms summarizing the results are shown in Figs. 5.15h to 5.15j. The lower bounds of the distributions provide the best indication of the resolution, as the segments include both single spots and spot clusters. The 5^{th}-percentile is 0.50 ± 0.13 µm in the x direction (along the illumination line). In the y direction (orthogonal to the illumination line i.e. the scan direction) a value of 0.53 ± 0.14 µm is found. The FWHM of the emission PSF according to the scalar diffraction based Airy distribution convolved with the pixel size is given by $0.57\lambda/\text{NA} = 0.43$ µm for an emission wavelength $\lambda = 563$ nm, and is included in Figs. 5.15h to 5.15j as black lines for comparison. The obtained values agree reasonably well with this theoretical value. The axial resolution was assessed in a similar fashion. For every segment, the maximum intensity in every layer $I_z(z)$ is determined. The values around the maximum are fitted with the parabolic

model $I_z(z) = w_z(z-z_c)^2 + I_{pz}$. Then, FWH$M_z = 2\sqrt{I_{pz}/2w_z}$. Figure 5.15j shows a histogram of the obtained result. The FWHM of the emission PSF is given by $1.61\lambda/(n-\sqrt{n^2-NA^2}) = 4.52\,\mu$m and is included as a black line. The 5^{th}-percentile of the found distribution is $3.35 \pm 0.10\,\mu$m, which exceeds the theoretical value. Factors impacting this analysis are the limited number of focal layers and the residual non-uniformity of the illumination intensity between the the focal layers.

5.5. Discussion

In summary, we presented a multi-line scanner for multi-focal fluorescence image acquisitions in a single scan. The sample is scanned with respect to a mechanically fixed optical setup, thereby providing simple means for an unlimited field of view. The optical architecture of the scanner is based on illumination PSF engineering. We used a design with diffractive optics for generating a set of parallel scan lines, focused at equidistant focal planes, in combination with a tilted sensor, for capturing the emitted fluorescence from the scan lines in parallel. An important new element in the design is the use of higher order astigmatism to improve the uniformity of peak intensity and line width along the scan lines. The approach is suitable for scanning large area tissue samples, as needed for imaging e.g. immunofluorescently labeled tissue samples. The ultimate goal for large area FISH imaging is to improve the statistical reliability in diagnosis, by increasing the number of cells tested and by correlating the resulting gene counts to the heterogeneous tissue structure.

The current setup is limited in the overall axial range of the 3D scan to a value between 10 and 15 µm. A larger axial range could for example be obtained by replacing the 20× objective with a 10× objective, leading to a fourfold increase in axial range. This, however, comes at the expense of lateral resolution, as the decrease in magnification is usually attended by a decrease in NA as well. Such an intended increase in axial range would also entail producing more than the current 9 parallel scan lines.

A major technical improvement of the current setup would be the use of an SLM with more pixels, e.g. 1920 × 1152 instead of the current 512 × 512. The enhanced space-bandwidth product could enable focal lines with an illumination NA matching the detection NA, equal to 0.75, instead of the current value 0.26, while maintaining or increasing the length of the scan lines of close to 0.5 mm. The main impact would be the improved sectioning capability, where our PSF simulations indicate that a peak in the background function with a width equal to 1.5 µm can be realized.

A next step for the proposed multi-line confocal scanner is the extension to multi-colour imaging. This can be realized using time multiplexing or space multiplexing. For time multiplexing the system cycles through the different excitation wavelengths. For each excitation wavelength a set of multi-focal scan lines must be generated, either via a fast switching SLM or by a setup with a modulator to select the wavelength and a dichroic beam splitter architecture wherein each color branch has a separate diffractive optical element, designed for the particular wavelength. For space multiplexing, the different sets of multi-focal scan lines for the different excitation wavelengths are slightly displaced with respect to each other in the direction of the tilted plane imaged onto the sensor. As a consequence, the

different color channels can be imaged in parallel, making more economical use of the sensor area. This approach does not require switching of the illumination and is compatible with a rolling shutter read-out.

Another extension that is envisioned is the application of techniques to improve the resolution. In particular, the Image Scanning Microscopy (ISM) method and the closely related pixel reassignment and rescan methods [42–45]. In fact, a modification of the TDI-approach we use can already deliver an ISM type of resolution improvement in the scan direction. This would require oversampling in the scan direction in combination with a digital stretch of the captured images in the scan direction before the step of adding the images acquired in subsequent frames. Although relatively straightforward, this would have the drawback of an anisotropic resolution, as the imaging in the field direction, along the scan lines, remains unaffected. A more isotropic improvement in resolution would require a modification of the multi-focal line illumination to a multi-focal multi-spot illumination, i.e. an illumination PSF which is spatially varying in the direction orthogonal to the scan direction as well. Such methods could overcome the drawback of the current scanning approach that it is not easily compatible with immersion objective lenses for high-resolution imaging.

Funding

Nederlandse Organisatie voor Wetenschappelijk Onderzoek (Netherlands Organization for Scientific Research - Applied and Engineering Sciences, NWO-TTW) (Project 13892).

Acknowledgments

We wish to thank Bas Hulsken, Rik Kneepkens, Jelte Peter Vink, Mathijs Rem, Gert-Jan van den Braak and Aslihan Arslan Carisey of Philips Digital & Computational Pathology, Hein Sleddens of the Erasmus University Medical Center and Hans Morreau of the Leiden University Medical Center for valuable research input.

References

[1] L. van der Graaff et al. “Multi-line fluorescence scanning microscope for multi-focal imaging with unlimited field of view”. In: *Biomedical Optics Express* 10.12 (2019), pp. 6313–6339. DOI: 10.1364/BOE.10.006313 (cit. on p. 99).

[2] F. Zanella, J. B. Lorens, and W. Link. “High content screening: seeing is believing”. In: *Trends in Biotechnology* 28.5 (2010), pp. 237–245. DOI: 10.1016/j.tibtech.2010.02.005 (cit. on p. 100).

[3] G. Bueno et al. “New Trends of Emerging Technologies in Digital Pathology”. In: *Pathobiology* 83.2-3 (2016), pp. 61–69. DOI: 10.1159/000443482 (cit. on p. 100).

[4] T. Dorval et al. “Filling the drug discovery gap: is high-content screening the missing link?” In: *Current Opinion in Pharmacology* 42 (2018), pp. 40–45. DOI: 10.1016/j.coph.2018.07.002 (cit. on p. 100).

[5] J. B. Pawley. *Handbook Of Biological Confocal Microscopy*. Ed. by J. B. Pawley. Springer US, 2006. ISBN: 978-0-387-45524-2. DOI: 10.1007/978-0-387-45524-2 (cit. on p. 100).

[6] L. Tian et al. “Computational illumination for high-speed in vitro Fourier ptychographic microscopy”. In: *Optica* 2.10 (2015), pp. 904–911. DOI: 10.1364/OPTICA.2.000904 (cit. on p. 100).

[7] E. Wang, C. M. Babbey, and K. W. Dunn. “Performance comparison between the high-speed Yokogawa spinning disc confocal system and single-point scanning confocal systems”. In: *Journal of Microscopy* 218.2 (2005), pp. 148–159. DOI: 10.1111/j.1365-2818.2005.01473.x (cit. on p. 100).

[8] T. Tanaami et al. “High-speed 1-frame/ms scanning confocal microscope with a microlens and Nipkow disks”. In: *Applied Optics* 41.22 (2002), pp. 4704–4708. DOI: 10.1364/AO.41.004704 (cit. on p. 100).

[9] H. L. Fu et al. “Optimization of a Widefield Structured Illumination Microscope for Non-Destructive Assessment and Quantification of Nuclear Features in Tumor Margins of a Primary Mouse Model of Sarcoma”. In: *PLoS ONE* 8.7 (2013). Ed. by J. P. Brody, e68868. DOI: 10.1371/journal.pone.0068868 (cit. on p. 100).

[10] T. C. Schlichenmeyer et al. “Video-rate structured illumination microscopy for high-throughput imaging of large tissue areas”. In: *Biomedical Optics Express* 5.2 (2014), pp. 366–377. DOI: 10.1364/BOE.5.000366 (cit. on p. 100).

[11] P.-H. Wang et al. “Non-axial-scanning multifocal confocal microscopy with multiplexed volume holographic gratings”. In: *Optics Letters* 42.2 (2017), pp. 346–349. DOI: 10.1364/OL.42.000346 (cit. on pp. 100, 101).

[12] A. Jesacher, S. Bernet, and M. Ritsch-Marte. “Colored point spread function engineering for parallel confocal microscopy”. In: *Optics Express* 24.24 (2016), pp. 27395–27402. DOI: 10.1364/OE.24.027395 (cit. on pp. 100, 101).

[13] L. Li et al. “High-throughput imaging: Focusing in on drug discovery in 3D”. In: *Methods* 96 (2016), pp. 97–102. DOI: 10.1016/j.ymeth.2015.11.013 (cit. on p. 100).

[14] R. Wolleschensky, B. Zimmermann, and M. Kempe. “High-speed confocal fluorescence imaging with a novel line scanning microscope”. In: *Journal of Biomedical Optics* 11.6 (2006), p. 064011. DOI: 10.1117/1.2402110 (cit. on p. 100).

[15] K.-b. Im et al. “Simple high-speed confocal line-scanning microscope”. In: *Optics Express* 13.13 (2005), pp. 5151–5156. DOI: 10.1364/OPEX.13.005151 (cit. on p. 100).

[16] C. Sheppard and X. Mao. “Confocal Microscopes with Slit Apertures”. In: *Journal of Modern Optics* 35.7 (1988), pp. 1169–1185. DOI: 10.1080/09500348814551251 (cit. on p. 100).

[17] C.-M. Chia et al. “Multiplexed holographic non-axial-scanning slit confocal fluorescence microscopy”. In: *Optics Express* 26.11 (2018), pp. 14288–14294. DOI: 10.1364/OE.26.014288 (cit. on p. 100).

[18] S. Abeytunge et al. “Rapid confocal imaging of large areas of excised tissue with strip mosaicing”. In: *Journal of Biomedical Optics* 16.5 (2011), p. 050504. DOI: 10.1117/1.3582335 (cit. on pp. 101, 111).

[19] S. M. Shakeri et al. “Optical quality assessment of whole slide imaging systems for digital pathology”. In: *Optics Express* 23.2 (2015), pp. 1319–1336. DOI: 10.1364/OE.23.001319 (cit. on pp. 101, 111).

[20] T. Yang et al. “Rapid imaging of large tissues using high-resolution stage-scanning microscopy”. In: *Biomedical Optics Express* 6.5 (2015), pp. 1867–1875. DOI: 10.1364/BOE.6.001867 (cit. on p. 101).

[21] P. J. Cutler et al. “Multi-Color Quantum Dot Tracking Using a High-Speed Hyperspectral Line-Scanning Microscope”. In: *PLoS ONE* 8.5 (2013). Ed. by D. Holowka, e64320. DOI: 10.1371/journal.pone.0064320 (cit. on p. 101).

[22] M. E. Gehm et al. “High-throughput, multiplexed pushbroom hyperspectral microscopy”. In: *Optics Express* 16.15 (2008), pp. 11032–11043. DOI: 10.1364/OE.16.011032 (cit. on p. 101).

[23] B. Hulsken. “Method for simultaneous capture of image data at multiple depths of a sample”. In: *US Patent No. 9,910,258 B2* (2018) (cit. on p. 101).

[24] J. Broeken, B. Rieger, and S. Stallinga. “Simultaneous measurement of position and color of single fluorescent emitters using diffractive optics”. In: *Optics Letters* 39.11 (2014), pp. 3352–3355. DOI: 10.1364/OL.39.003352 (cit. on p. 101).

[25] A. Greengard, Y. Y. Schechner, and R. Piestun. “Depth from diffracted rotation”. In: *Optics Letters* 31.2 (2006), pp. 181–183. DOI: 10.1364/OL.31.000181 (cit. on p. 101).

[26] Y. Shechtman et al. “Optimal Point Spread Function Design for 3D Imaging”. In: *Physical Review Letters* 113.13 (2014), p. 133902. DOI: 10.1103/PhysRevLett.113.133902 (cit. on p. 101).

[27] Y. Shechtman et al. “Multicolour localization microscopy by point-spread-function engineering”. In: *Nature Photonics* 10.9 (2016), pp. 590–594. DOI: 10.1038/nphoton.2016.137 (cit. on p. 101).

[28] M. Siemons et al. “High precision wavefront control in point spread function engineering for single emitter localization”. In: *Optics Express* 26.7 (2018), pp. 8397–8416. DOI: 10.1364/OE.26.008397 (cit. on pp. 101, 116).

[29] C. Smith et al. “Simultaneous measurement of emission color and 3D position of single molecules”. In: *Optics Express* 24.5 (2016), pp. 4996–5013. DOI: 10.1364/OE.24.004996 (cit. on pp. 101, 102).

[30] B. C. Chen et al. “Lattice light-sheet microscopy: Imaging molecules to embryos at high spatiotemporal resolution”. In: *Science* 346.6208 (2014). DOI: 10.1126/science.1257998 (cit. on p. 101).

[31] B.-J. Chang et al. “Universal light-sheet generation with field synthesis”. In: *Nature Methods* 16.3 (2019), pp. 235–238. DOI: 10.1038/s41592-019-0327-9 (cit. on p. 101).

[32] B. R. W. Gerchberg and W. O. Saxton. “A Practical Algorithm for the Determination of Phase from Image and Diffraction Plane Pictures”. In: *Optik* 35.2 (1972), pp. 237–246. DOI: 00.0000/lg.gerchberg1972 (cit. on p. 103).

[33] S. Stallinga. “Finite conjugate spherical aberration compensation in high numerical-aperture optical disc readout”. In: *Applied Optics* 44.34 (2005), pp. 7307–7312. DOI: 10.1364/AO.44.007307 (cit. on p. 109).

[34] S. M. Shakeri. “Whole Slide Imaging Systems for Digital Pathology”. Dissertation. 2017. ISBN: 978-94-6186-781-0. DOI: 10.4233/uuid:4890151d-0052-4107-93f1-3ab2c5283cf8 (cit. on p. 109).

[35] H. Netten et al. “A fast scanner for fluorescence microscopy using a 2-D CCD and time delayed integration”. In: *Bioimaging* 2.4 (1994), pp. 184–192. DOI: 10.1002/1361-6374(199412)2:4%3C184::AID-BIO3%3E3.0.CO;2-M (cit. on p. 112).

[36] E. Ronzitti et al. “LCoS nematic SLM characterization and modeling for diffraction efficiency optimization, zero and ghost orders suppression”. In: *Optics Express* 20.16 (2012), pp. 17843–17855. DOI: 10.1364/OE.20.017843 (cit. on p. 113).

[37] M. J. Booth. “Wavefront sensorless adaptive optics for large aberrations”. In: *Optics Letters* 32.1 (2007), pp. 5–7. DOI: 10.1364/OL.32.000005 (cit. on p. 117).

[38] F. de Chaumont et al. “Icy: an open bioimage informatics platform for extended reproducible research”. In: *Nature Methods* 9.7 (2012), pp. 690–696. DOI: 10.1038/nmeth.2075 (cit. on pp. 122, 123).

[39] I. Khaw et al. “Flat-field illumination for quantitative fluorescence imaging”. In: *Optics Express* 26.12 (2018), pp. 15276–15288. DOI: 10.1364/OE.26.015276 (cit. on p. 123).

[40] L. van der Graaff et al. “Fluorescence imaging for whole slide scanning using LED-based color sequential illumination”. In: *Proc. SPIE 10679, 106790D* (2018). DOI: 10.1117/12.2306776 (cit. on p. 123).

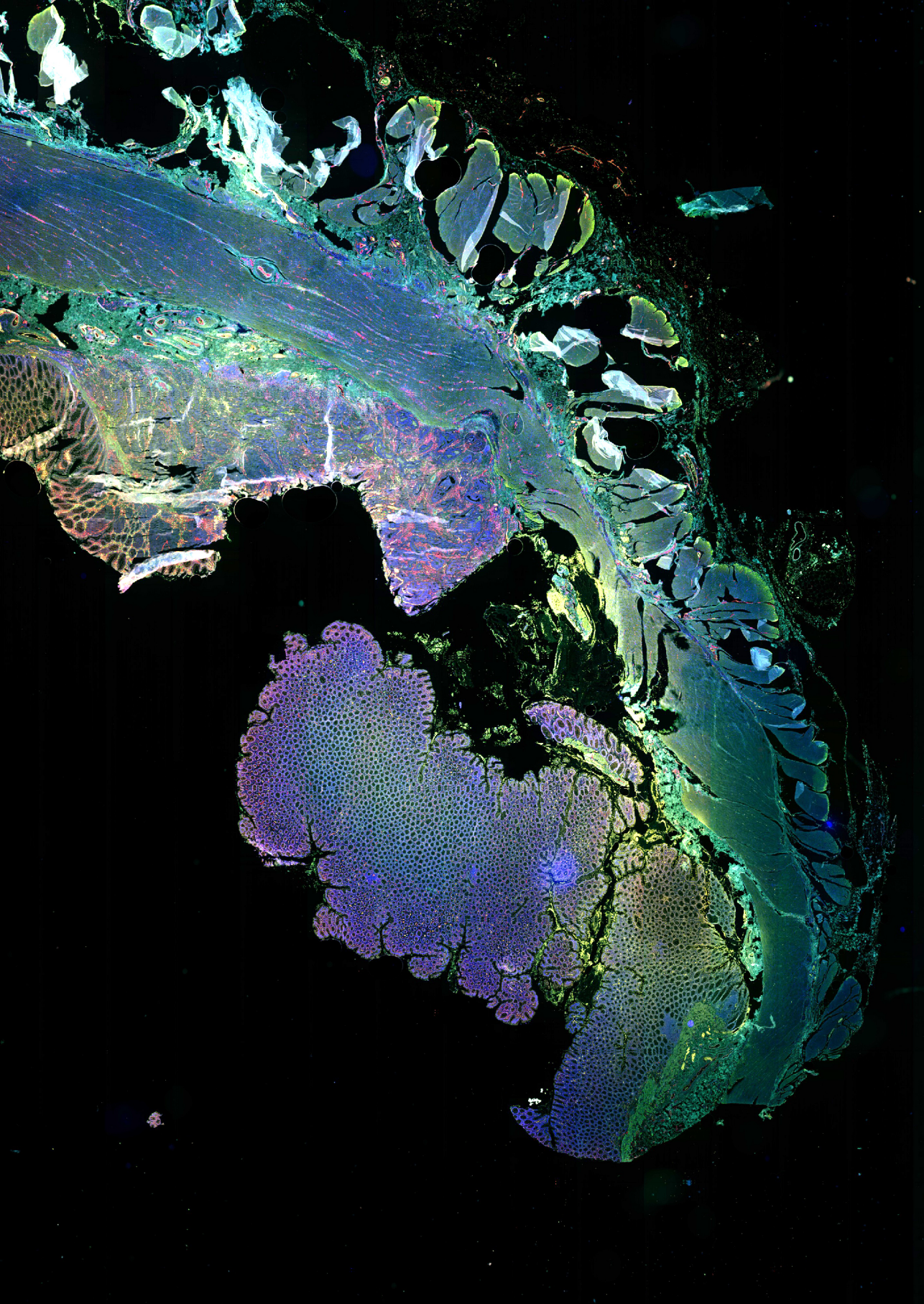
[41] I. Smal et al. “Quantitative Comparison of Spot Detection Methods in Fluorescence Microscopy”. In: *IEEE Transactions on Medical Imaging* 29.2 (2010), pp. 282–301. DOI: 10.1109/TMI.2009.2025127 (cit. on p. 125).

[42] C. B. Müller and J. Enderlein. “Image Scanning Microscopy”. In: *Physical Review Letters* 104.19 (2010), p. 198101. DOI: 10.1103/PhysRevLett.104.198101 (cit. on p. 127).

[43] A. G. York et al. “Instant super-resolution imaging in live cells and embryos via analog image processing”. In: *Nature Methods* 10.11 (2013), pp. 1122–1126. DOI: 10.1038/nmeth.2687 (cit. on p. 127).

[44] S. Roth et al. “Optical photon reassignment microscopy (OPRA)”. In: *Optical Nanoscopy* 2.1 (2013), p. 5. DOI: 10.1186/2192-2853-2-5 (cit. on p. 127).

[45] G. M. De Luca et al. “Re-scan confocal microscopy: scanning twice for better resolution”. In: *Biomedical Optics Express* 4.11 (2013), pp. 2644–2656. DOI: 10.1364/BOE.4.002644 (cit. on p. 127).



6

Discussion and Outlook

New imaging modalities for Whole Slide Imaging (WSI) systems were investigated in this thesis. A novel optical architecture based on a tilted multi-line imaged sensor was explored. This platform makes multi-focal WSI available, and was used to demonstrate quantitative phase tomography (QPT) and imaging of thick tissue sections with enhanced contrast. Two modalities for fluorescence imaging on a WSI system were designed and built: widefield fluorescence with an LED illumination source and multi-focal multi-line scanning with a laser source. Here, we summarize the results of this work and present an outlook on future research directions.

6.1. Discussion

The work presented in this thesis was carried out in the framework of a publicly funded project (NWO-TTW, Nederlandse Organisatie voor Wetenschappelijk Onderzoek - Technische en Toegepaste Wetenschappen) with private support of Philips Digital & Computational Pathology. The goal was to explore fluorescence and other use cases of a new whole slide scanner platform, based on the use of a multi-line image sensor that is placed at an oblique angle with the optical axis. The goal of multi-line image acquisition with this tilted sensor is to make closed-loop autofocus scanning possible of conventional thin tissue sections. A novel CMOS multi-line sensor has been devised by Philips for this purpose [1–3].

In chapter Chapter 2 a system characterization of this sensor is presented. The sensor was characterized for gain and noise, and a system model was developed to find the optimal SNR given the available photo-electron flux. The gain in the default mode of operation was measured to be 50 e^-/ADU. In the limit of a low photon flux the optimal gain was found to be 32 e^-/ADU. The system model showed that images with a very high SNR of 292 can be acquired, provided that a sufficiently high photon flux can be realized. This implies that 8-bit images can have a noise level less than 1 LSB. For the anticipated low light conditions in fluorescence imaging, however, the levels of read noise and dark current make this sensor unsuitable for generating high quality images. Two major issues with the sensor were found with our experimental characterization. At high line rates, the sensor showed missing symbols, leading to non-linearities in the read out. The discontinuity in the readout dynamic range has a size of about 200 e^-. The sensor also showed high frequent fluctuations in the gain, leading to an apparent increase in the rms noise levels up to 12 %.

The WSI-architecture with the tilted multi-line sensor can also function as a platform for 3D WSI-imaging, which is explored in Chapter 3. Multi-focal '3D' imaging was demonstrated by acquiring 8 layers in a single scan pass. The obtained volumetric image data was used to test two novel contrast modalities based on computational imaging. The first modality is a deconvolution technique for improving the inherently low axial contrast of multi-focal images of thick specimens. The deconvolution algorithm is a variant of the Iterative Constrained Tikhonov-Miller (ICTM) deconvolution method, targeting improvement of axial contrast, rather than in-plane resolution. The technique is demonstrated on scans of 60 µm thick tissue slides. The second modality is 3D phase imaging based on the recently introduced QPT method [4]. Comparison of QPT images of a slide prepared for FISH to the fluorescence reference image reveals that QPT imaging is able to image the sites labeled for Fluorescence *in situ* Hybridization (FISH) imaging. QPT imaging of an unstained tissue slide and comparison to immunofluorescence imaging shows that phase imaging can provide additional structural information on tissues. The image formation theory of QPT points to an in-plane transfer function that has the character of a band-pass spatial frequency filter. A system design study revealed that the filter characteristic is entirely given by the diffraction length scale, the scaled imaging NA and the number of scanned layers. This study might be used as a generic design guide for implementation of QPT on a multi-focal brightfield system. For both considered techniques simplified algorithms were developed compatible with parallel processing at very high speeds. A processing through-

put exceeding the image acquisition throughput was realized, which indicates that these algorithms are suitable for on-line, and potentially on-chip, processing.

Chapter 4 describes the development of a widefield multi-color fluorescence imaging scanner. The system uses an epi-illumination system that is based on LEDs in combination with multi-band dichroics. A color sequential illumination strategy was used for multi-color imaging using a single monochromatic sensor. This solution is technologically robust, cost effective, and requires the addition of just a few components to a brightfield scanning system. The photo-electron yield was measured for three commonly used fluorophores, and the system was found to be able to capture between roughly 200 and 1000 $e^-/\mu m^2/ms$. Further, it was determined that about $10^5 e^-/\mu m^2$ are required to obtain a Fourier Ring Correlation (FRC) resolution close to the diffraction limit. The substantial exposure time to capture this amount of photo-electrons forms the main speed limitation in fluorescence imaging with LED illumination. Intrinsically, a step-and-stitch approach will therefore be faster than line scanning for fluorescence imaging with an incoherent illumination source because it makes better use of the illumination étendue. This disadvantage of line scanning is partly overcome by the use of TDI. A system with 96 TDI lines was estimated to achieve a reasonable throughput of about 130 kPixel/s. This makes scanning possible of an area of $15 \times 15\,mm^2$ in three colors in about 23 min, which might be sufficient for clinical applications where occasional scanning of fluorescent slides is needed.

In Chapter 5 a system was proposed for multi-focal fluorescence image acquisitions. A laser illumination source was used for high power illumination that overcomes the étendue limitations of the LED-based fluorescence WSI system. Illumination PSF engineering using diffractive optics was applied to generate a set of parallel scan lines in object space, that span a plane conjugate to a tilted detector. This system thus builds on the concept of multi-focal scanning with a tilted sensor. An important new element in the design was the use of higher order astigmatism to improve the uniformity of peak intensity and line width along the scan lines. The prototype system was able to acquire 8 focal layers simultaneously at a close to diffraction limited resolution. Focusing the illumination on the sample provided a confocal suppression of background, albeit limited due to the relatively low illumination NA. A peak Signal-to-Noise-Ratio (pSNR) of 34 was realized in scanning a sample labeled using FISH. A throughput of 0.29 MPixel/s was realized in the prototype system. Corrected for global shutter constraints, non fundamental inefficiencies, and the relatively low laser power that was used, it was concluded that this optical architecture can achieve a throughput of several hundreds MPixel/s. This would enable scanning an area of $15 \times 15\,mm^2$ in 8 layers in less than a minute, which is on the same order of magnitude as is achieved in brightfield WSI systems.

6.2. Outlook

In this final section an outlook to opportunities for future research in WSI systems is provided. First, opportunities for multi-color imaging are given. Then, we discuss the challenges and some potential solutions for using a general purpose scientific CMOS (sCMOS) camera for multi-line scanning of a tilted object plane. Further, some opportunities for extension of WSI techniques to Light Sheet Microscopy (LSM) and Structured Illumination Microscopy (SIM) are considered. Fi-

nally, we discuss the opportunities for computational imaging techniques for WSI.

6.2.1. Multi-color imaging opportunities

A next step for the proposed multi-line confocal scanning system presented in Chapter 5 might be the extension to multi-color imaging. A color sequential illumination approach can be used for time multiplexing the color images, in a similar fashion as was used for the LED based widefield system. The sample is illuminated sequentially with light of different wavelengths, targeting different fluorophores in the sample. Only a single monochromatic sensor is required then, and the color channels are therefore intrinsically aligned. This approach increases the scan time with a factor equal to the number of color channels. Alternatively, a spatial multiplexing approach might be used, by giving the illumination lines a slight displacement with respect to each other, see Fig. 6.1a. This approach realizes a more economical use of the sensor real estate. It makes the retrieval of a multi-color frame in a single sensor readout possible without increasing the scan time, provided that enough electronic bandwidth is available. In addition, it does not require fast synchronized switching of the illumination source.

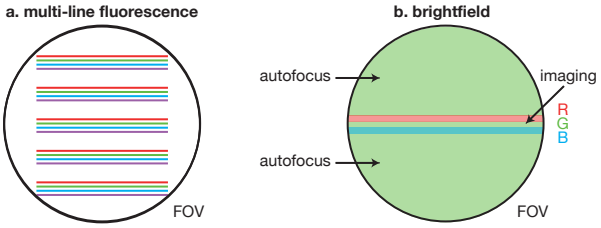


Figure 6.1: Illustration of spatial multiplexing for multi color imaging. (a) Arrangement of a four color illumination line pattern in the field of view for multi-line imaging. (b) Arrangement for spatial multiplexing for 2D imaging with autofocus. The central three rows are used for RGB color imaging. The rest of the field of view is illuminated with green light to obtain autofocus information.

Spatial multiplexing also forms an opportunity for an increased throughput in 2D brightfield imaging with continuous autofocus in the scanning system devised by Philips, see Chapter 3. Simultaneous acquisition of the red, green and blue channel might result in a threefold increase in throughput at the same SNR and without the need for shorter integration times, provided, again, that electronic bandwidth is not a bottleneck. A potential implementation is illustrated in Fig. 6.1b, where the center of the FOV is illuminated by three rows of red, green and blue light. The illumination rows are aligned with the sensorlets on the sensor, such that every row matches a separate sensorlet. Now, instead of using only the center sensorlet for imaging, three sensorlets could be read out simultaneously and combined in digital post-processing to obtain an RGB color image. This readout scheme can potentially be realized within the current capabilities of the sensor's on-board FPGA and would not need any adjustments to the sensor. The rest of the FOV can be illuminated monochromatically, e.g. with green light for creating an autofocus signal.

Spatial multiplexing of the color channels does require an adjustment to the il-

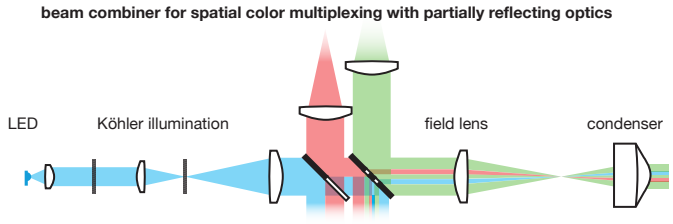


Figure 6.2: Schematic outline of an illumination unit to create a spatially multiplexed multi-color illumination. The partially reflecting mirrors are illustrated with the reflecting parts in black and the transparent parts in white. For the blue channel the LED light source and Köhler illumination are drawn. The other channels have the same optical design which is omitted in this illustration for the sake of clarity.

lumination source. Figure 6.2 illustrates a potential implementation where a beam combiner based on partially reflective optics (patterned mirrors) is used to create the desired chromatic patterning of the illumination FOV. The field at the last mirror is imaged onto the object plane to create the desired illumination pattern. The image of the mirror will give a tilted image in object space. This is not a significant problem, as the red, green, and blue lines are imaged within the focal depth. Possible defocus might even be minimized by matching the tilt of the mirror with the sensor tilt. This design does require individual flat fielding correction for each color channel. The use of partially reflective optics removes the need for dichroic mirrors in the illumination unit. Multi-focal imaging with spatially multiplexed color illumination would require an illumination pattern of repeated red, green, and blue rows aligned with the sensorlets.

It is interesting to speculate on the possibilities of QPT in the context of multi-color imaging. The local phase is essentially the ratio of the local refractive index and the wavelength. Measuring the refractive index for multiple wavelengths, e.g. using the red, green, and blue channel of an RGB system, thus provides information on the refractive index as a function of wavelength. The local dispersion can then be measured by numerically fitting e.g. the Cauchy-relation between refractive index and wavelength. This would enable quantitative dispersion microscopy [5]. The dependence of the Optical Transfer Function (OTF) on the wavelength forms a challenge to bring this idea to practice.

6.2.2. Multi-line scanning with a general purpose sCMOS sensor

The proposed confocal scanning method in Chapter 5 was devised to be compatible with the optical architecture of Philips' multi-line sensor based WSI platform, see Chapter 2. This sensor is able to read out up to 8 lines simultaneously and a very high fill factor is realized by placing the read out electronics between the light sensitive pixel [2]. Off-the-shelf sCMOs sensors, however, provide the advantage of having a very low read noise such as is required for fluorescence imaging. Efficient implementation using a general purpose sCMOS sensor requires that two challenges are addressed. In the proposed design we considered a global shutter read-out of the sensor, where all rows on the sensor have a synchronized integration time. Most sCMOS camera's, however, read out the whole sensor row-by-row, combined with a rolling shutter read-out to minimize the dead time [6]. A rolling

shutter read-out is compatible with the proposed multi-line scanning system, although two minor problems have to be addressed. First, it should be possible to select multiple regions of interest for read-out, while skipping the lines in between. Although there is no physical limitation for skipping lines in the read out, this option is generally not available as a standard option and would require custom firmware. Second, the rolling shutter introduces a misalignment of up to a pixel in between the first and last scan line, corresponding to the top and bottom layer of the scanned multi-focal volume. This can be incorporated in the lateral alignment of the scanned focal layers in digital post-processing. A rolling shutter read-out is fully compatible with spatially multiplexed multi-color imaging, but can not be combined with time multiplexed multi-color imaging without introducing significant dead time, as the light sources need to be synchronized with a global shutter.

The second challenge is that sCMOS camera's are not designed to be used for imaging planes at an oblique angle with the optical axis. The sensors are usually equipped with an on-chip microlens array for an increased effective fill factor [7], which entails that light incident under an angle with the normal to the sensor plane is laterally displaced upon incidence at the silicon. This gives rise to the unwanted side effect of a significantly reduced photon detection sensitivity. For example, the camera that was used in Chapter 5 showed a reduction in sensitivity of 53 % for light incident under a 20° angle. Another minor disadvantage is that placing the sensor under an angle results in anisotropy of the pixel grid projected back to the object plane. The back-projected pixel size along the scan axis becomes slightly smaller than along the orthogonal field axis, which has to be corrected for in digital post-processing.

6.2.3. Optical methods for creating a tilted object plane

A potential solution for the issue of oblique incidence on a general purpose sCMOS sensor with a microlens array is to create the tilted object plane by means of an additional set of relay lenses, instead of tilting the sensor. Than can be done similarly to optical architectures used in Oblique Plane Microscopy (OPM) [8] or Swept, Confocally Aligned Planar Excitation (SCAPE) microscopy [9]. In both techniques, a $4F$ system consisting of an objective lens and tube lens are used to image a strongly tilted ($\sim 60°$) object plane to an even further tilted image plane. Then, the light is refocused by a second $4F$ system consisting of an objective and tube lens. Finally, a third $4F$ system consisting of an objective and tube lens is placed perpendicular to the tilted image plane to project a conjugate plane onto the sensor. As a result, the light is perpendicularly incident on the sensor while the relay path creates a tilted image plane. This method is fully translatable to the proposed method of multi-line confocal imaging. A further advantage of this approach is that it does not require the sensor to be designed for a specific tilt angle. Additionally, this would enable covering a significantly larger axial range than in the current setup.

We propose here a much simplified version of the optical architecture of OPM and SCAPE. In our proposal a single relay system, consisting of two lenses in a telecentric $4F$ configuration, is placed after the main optical imaging system comprising the objective and tube lens. This relay system images the intermediate image plane after the tube lens onto the image sensor such that the chief rays of the

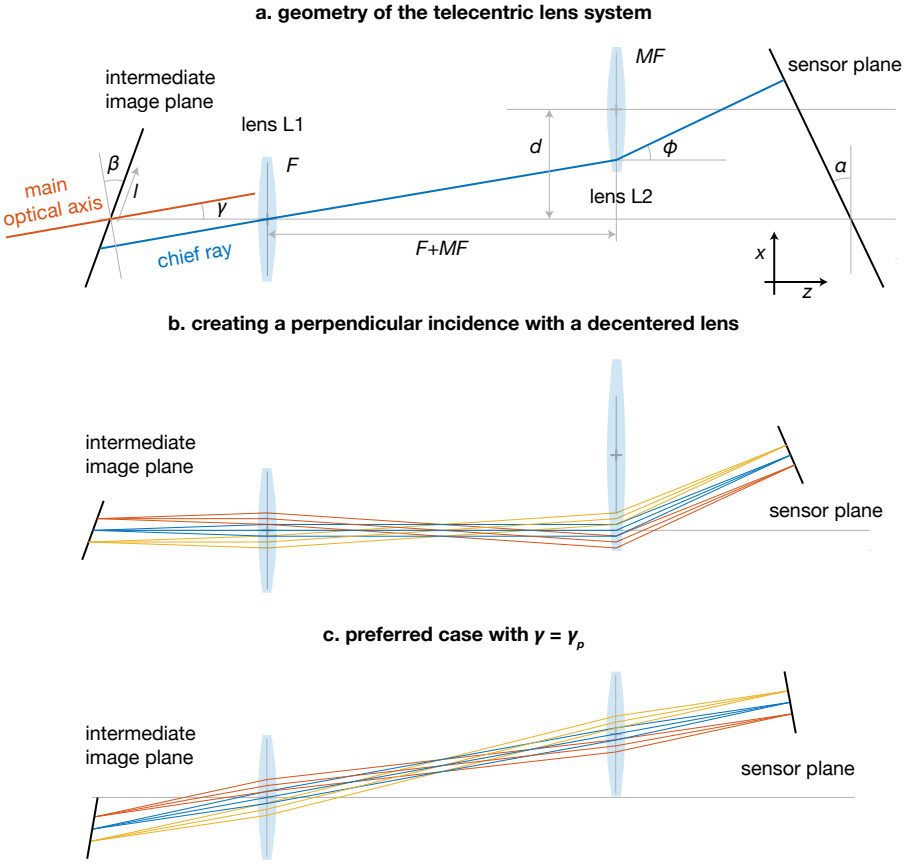


Figure 6.3: Illustration of the use of a telecentric relay lens system with an intentional decenter for creating perpendicular incidence on the sensor plane. (a) The proposed geometry. (b) Example of creating perpendicular incidence at the sensor plane using decentered lenses. (c) Illustration of the preferred case, where the telecentric relay lens system is placed obliquely with respect to the main optical axis of the scanner to minimize the field angles and aperture height within the relay system. In both examples, we use $M = 1$, F =50 mm, β =–20 ° and NA =.034 in the intermediate image plane, matching the imaging system described in Chapter 5.

beams that are incident on the image sensor are perpendicular to the sensor plane. This simplified relay system is possible because the object plane in sample space makes a relatively small angle of about 0.9° with the standard focal plane normal to the optical axis of the objective lens, much smaller than in OPM or in SCAPE microscopy. The key principle of the proposed approach is an intentional decenter between the two lenses of the relay system. Figure 6.3b shows an example of this principle. The relative decenter of the lenses results in a change of the angle of incidence of the rays towards the image points on the sensor. This give the freedom of design to create perpendicular incidence on the sensor while still having a tilted object plane.

We will now elaborate on the thin lens design for such a tilting $4F$ relay system. The design goal is to have perpendicular incidence of the chief rays originating from the tilted intermediate image plane onto the sensor plane. The tilt angles of the sensor plane and the angle of the chief rays in the exit pupil as a function of the lens parameters will be derived and solved for this condition of perpendicular incidence. The geometry of the relay lens system is outlined in Fig. 6.3a. Two lenses L1 and L2 are used with focal lengths $F_1 = F$ and $F_2 = MF$, where M is the magnification of the relay system, and the lenses have a spacing $F + MF$, which makes the relay system nominally telecentric. An intentional decenter d is introduced between the optical axes of L1 and L2. Geometrical optics implies that an image point P depends on the object coordinates (z_o, x_o) by:

$$(z_i, x_i) = P(z_o, x_o) = \left(M^2 z_o \,,\; -M x_o + \frac{M d z_o}{F} \right). \tag{6.1}$$

The coordinates (z_o, x_o) are taken with respect to the front focal point of L1 and the coordinates (z_i, x_i) are taken with respect to the back focal point of L2. The coordinate axes are chosen such that the z-axis is perpendicular to the common principal plane of the two lenses. The system is telecentric and therefore the angle θ_i under which a ray leaves the relay system only depends on the angle of incidence θ_o. From Eq. (6.1), it can be derived that this relation is given by

$$\theta_i = R(\theta_o) = \arctan\left(\frac{d}{FM} - \frac{1}{M} \tan\theta_o \right). \tag{6.2}$$

It turns out to be convenient to utilize a second degree of freedom. Namely, the relay system can be oriented at an angle γ with the main optical axis of the scanner, defined by the objective and tube lens. The intermediate image plane after the tube lens of the scanner, which is thus the object plane of the relay system, can then be described by the coordinates in the object space of the relay system as:

$$(z_p, x_p) = \big(-l \sin(\beta + \gamma) + z_{p0} \,,\; l \cos(\beta + \gamma) + x_{p0} \big), \tag{6.3}$$

where β is the tilt angle of the intermediate image plane with respect to the main optical axis of the scanner, l is the coordinate along the intermediate image plane and x_{p0} and z_{p0} are arbitrary offsets that turn out not to affect the outcome of this analysis. The sensor is placed at the image of this intermediate image plane formed by the relay system, with the coordinates $(z_s, x_s) = P(z_p, x_p)$ obtained from Eq. (6.1). The tilt angle α of the sensor plane with respect to the x-axis then follows

as:

$$\tan\alpha = -\frac{\mathrm{d}z_s}{\mathrm{d}x_s} = -\left(\frac{FM\sin\left(\beta+\gamma\right)}{F\cos\left(\beta+\gamma\right)+d\sin\left(\beta+\gamma\right)}\right). \tag{6.4}$$

The chief rays entering the lens system are parallel to the main optical axis and have therefore an angle of incidence γ onto the relay lens system. The angle of the chief rays in the exit pupil then follows as $\phi = R(\gamma)$ according to Eq. (6.2). For the desired perpendicular incidence at the sensor we must then solve the requirement $\tan\phi = \tan\alpha$ for obtaining an expression for the relative displacement d of the lenses of the relay system.

The preferred case is shown in Fig. 6.3c. In this preferred case the main optical axis of the imaging system passes through both the principal points of L1 and L2. This configuration minimizes the required field angles and aperture heights. The value of γ is now equal to $\gamma_p = \arctan\left(d/(F+MF)\right)$. Solving d for perpendicular incidence of the chief rays at the sensor plane now results in the simple relation:

$$d_p = -FM\tan(\beta), \tag{6.5}$$

which immediately gives:

$$\tan\gamma_p = \frac{M}{M+1}\tan\beta. \tag{6.6}$$

The values of d_p and γ_p as a function of the intermediate image plane tilt are plotted for different values of the relay magnification M in Fig. 6.4. For a unit magnification, reasonably small field angles (scaling with γ_p) are observed, e.g. for $\beta = 20°$ a value of $\gamma_p = 10.3°$ is found. Lower magnifications lead to lower values of γ_p, in the limit of high magnifications $\gamma_p \to \beta$. Using a negative focal length for lens L1 might be a way to create a compact lens design. at expense of higher field angles. However, the value of γ_p diverges to unfeasible large values of when approaching $M = -1$, see Eq. (6.6). A final, more accurate, optical design should incorporate the effects of finite lens thicknesses and large field angles found here, which are effects beyond the realm of paraxial geometrical optics.

6.2.4. Scanning Structured Illumination Microscopy

Another extension that is envisioned for the laser based fluorescence imaging system is the application of techniques to improve the resolution. This would provide the advantage of making scanners with air-incidence objectives (e.g. NA = 0.75) that could be made competitive in resolution to scanners with immersion objectives (e.g. NA = 1.3 to 1.5) without having to bother with immersion liquids in an automated scanner environment. The proposed imaging system samples the emission PSF in the scan direction, opening the possibility of enhancing the resolution by using the Image Scanning Microscopy (ISM) method or the closely related pixel reassignment and rescan methods [10–13]. Although relatively straightforward, this would have the drawback of an anisotropic improvement of resolution, as the imaging in the field direction orthogonal to the scan direction remains unaffected.

A more isotropic improvement in resolution would require a modification of the line illumination to a multi-spot illumination. This would enable a true Structured Illumination Microscopy (SIM) approach for resolution enhancement. In the

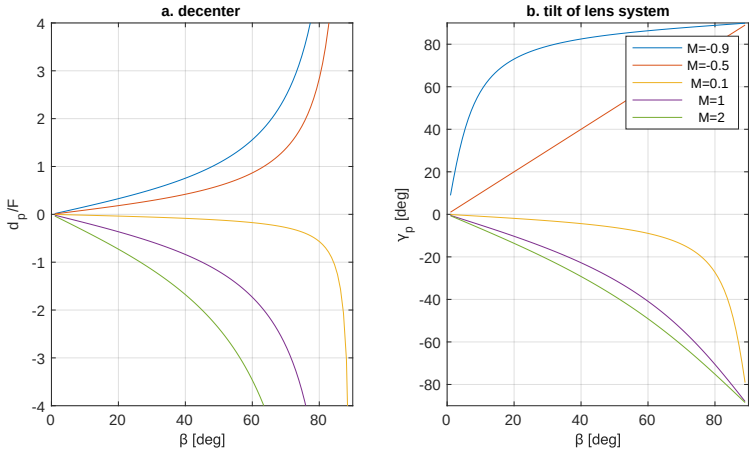


Figure 6.4: Geometry of the telecentric lens system in the preferred case of $\gamma = \gamma_p$ as a function of the tilt of the intermediate image plane β for different magnifications M. (a) The decenter d_p. (b) The tilt of the lens system with respect to the main optical axis γ_p.

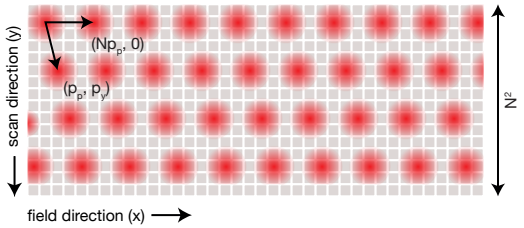


Figure 6.5: Illustration of a spot grid array illumination pattern for scanning SIM. The sensors pixel grid is illustrated in the gray background. The spots have a pitch along the field direction of Np, where p is equal to sensors pixel pitch and $N = 4$. The grid consist of rows of spots that are mutually shifted over p. The most densely packed situation is illustrated where the row spacing equals Np, such that a region of interest of N^2 consecutive rows is used to acquire the data for a resolution enhanced image.

technique of SIM [14–17] the object is illuminated with a spatially periodic light pattern consisting of bright and dark fringes and a series of images is recorded for different translations and rotations of the illumination pattern with respect to the object. An image with up to two times better spatial resolution can be obtained from the set of raw images by an image reconstruction and filtering operation [18].

This scheme can be adapted to a WSI system by making two major modifications. The first is the change of a line pattern to a dot pattern with period Np along the line sensor. The second modification is the means of displacement of the illumination pattern with respect to the object. Instead of translating and rotating the line position, the scanning movement of the sample with respect to the illumination pattern is exploited, such that an illumination pattern can be used that is static with respect to the scanner. By reading out the N (typically 4) lines on the sensor surrounding a single row of illumination spots, data is acquired for N spot positions shifted over a range Np_s in the scan (y) direction with respect to the object, where p_s is the pitch of the stage steps (typically 1 px). The scanning

is in one direction only, which leaves the displacement in the field (x) direction unaccounted for. This is solved by imaging the same position on the sample multiple times with a set of N illumination rows, such that the illumination pattern in adjacent rows is shifted over the pattern pitch p_p. The illumination pattern thus has the shape of a spot array with a parallelogram-shaped unit-cell with the basis vectors of the corresponding Bravais lattice equal to $(Np_p, 0)$ and (p_p, p_y), where $(0, 1$ is in the scan (y) direction, see Fig. 6.5. As a result, every pixel sized area in the sample is measured N^2 times, corresponding to the excitation with a spot shifted over (lp_p, mp_s) with $l, m \in 1...N$. By taking linear combinations of (the Fourier transforms of) these measurements, the conventional spatial frequency band images can be obtained. These images can be further filtered and combined into the final SIM image with existing image reconstruction techniques [18]. The pitch p_p and p_s should approach $\lambda/2\text{NA}$ for significant resolution improvement and should be equal for isotropic sampling. An optimum use of sensor real estate is obtained by choosing $p_p = p_s = 1\,\text{px}$ equal to the pixel pitch and $p_y = N\text{px}$, as is illustrated in Fig. 6.5. In this case data can be acquired with the minimum of N^2 consecutive rows on the sensor.

The dot pattern can be made with four beam interference of laser light using two (nearly) crossed gratings placed in a position in the illumination light path conjugate to the objective lens back focal plane, or by means of a phase pattern created by a spatial light modulator (SLM). A final design should incorporate polarization effects, which have impact on the modulation depth and precise multi-spot pattern shape.

A generalization to 3D dot patterns by mixing in the central 0th order diffracted beam can be considered as well. This 0th order beam creates an interference pattern that depends on the axial position as well, which is needed for making an improvement in axial resolution and contrast (optical sectioning) [19]. A slight repositioning of the illumination spots in the back focal plane of the objective can be used to tilt the 3D illumination pattern to match the tilted sensor plane. For limited tilt angles, no significant defocus will be introduced within the N^2 rows used to modulate the illumination in the lateral directions. This could possibly enable 3D SIM in a single scan movement with an illumination pattern that is static with respect to the scanner.

6.2.5. Scanning Light Sheet Microscopy

Light sheet microscopy (LSM) is another imaging technique that might benefit from a stage scanning approach for WSI with an unlimited FOV. LSM is a wide-field fluorescence microscopy technique in which optical sectioning is realized by illuminating only a thin 'sheet' within the sample, i.e. the illumination light propagates in the object plane, in contrast to epi-illumination or brightfield illumination where the light propagates through the object plane [20–22]. In one of the earliest implementations of LSM [20], the light sheet is focused into the sample by means of a second objective lens that is placed at an angle of 90° with respect to the optical axis of the main imaging objective lens. This design comes with strong mechanical constraints on the geometry of the object of interest and makes it incompatible with the use of glass microscopy slides. This is solved in OPM [8], where a single high NA objective lens is used for both illumination and detection. A part of

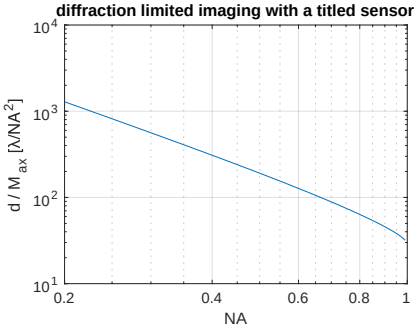


Figure 6.6: The axial range in object space that can be imaged with a tilted sensor while remaining diffraction limited over the whole range ($W_{\mathrm{rms}} < 72\,\mathrm{m}\lambda$). The axial range is given in units of the diffraction length scale λ/NA^2. The range is independent of the magnification and sensor tilt. A sample refractive index of $n = 1.5$ and an air objective are assumed.

the aperture is used to create a strongly inclined illumination sheet, while another (limited) part of the aperture is used for detecting the fluorescence signal. As discussed earlier in this text, a relay path consisting of two additional microscope systems is used in OPM to image the tilted object plane onto a sensor. These two microscopes are tilted with respect to each other such that the image is formed by beams that are perpendicularly incident on the sensor. This design is further extended in SCAPE microscopy [9], where the light sheet is scanned through the sample by means of a scanning mirror, which results in a high frame rate volumetric imaging modality. These principles can be combined with a stage scanning approach in a relatively straightforward way, providing an extension of LSM to the realm of WSI. The OPM optical layout can be used for this purpose as it enables to make an image of an oblique sample plane illuminated with a light sheet. The optical design can even be simplified as there is no need for scanning the illumination through the imaging volume. Instead, stage scanning in combination with the tilted object plane enables multi-focal imaging with unlimited FOV by a single scan movement.

Another inroad in which the current thesis work might impact LSM is in the use of a tilted image sensor for imaging a tilted object plane illuminated by a light sheet. This would remove the optical complexity of a relay path consisting of two oblique positioned $4F$ systems in LSM techniques such as OPM and SCAPE. In general, this is prohibited by the spherical aberration that is introduced by imaging at multiple focus depths simultaneously [8, 23]. However, in the work presented in Chapter 3 we found that it is possible to significantly reduce this spherical aberration by adjusting the free working distance and the axial position of the sensor. In Eq. (3.12) an expression is provided for the total RMS aberration W_{rms} for a finite conjugate imaging system in focus, as a function of the axial position of a row of pixels on the sensor z_3, the total axial range in image space d and the axial position of the center of the sensor z_c. An optimum center position of the sensor was derived and is provided in Eq. (3.13). For this optimum sensor position, W_{rms} has its minimum at the center of the axial range $z_3 = z_c$ and increases symmetrically for off-center positions. The maximum aberration is therefore found at the edge of the axial range, at $z_3 = z_c \pm d/2$. Substituting these values for z_3 and z_c in the expression for W_{rms} given in Eq. (3.12) provides the maximum aberration as a function of

the axial range in image space d. This equation has the shape:

$$W_{\mathrm{rms}} = \xi \frac{d}{M_{\mathrm{ax}}}, \tag{6.7}$$

where M_{ax} is the axial magnification and ξ a dimensionless number which only depends on the NA and the refractive indices of the sample and the medium between the objective and the cover slip (air or immersion medium). This number can be shown to be:

$$\xi^2 = -\frac{g_{1,1}\left(4\,g_{1,2}{}^2\,g_{1,3}{}^2\,g_{3,3} - g_{1,1}\,g_{1,2}{}^2\,g_{3,3}{}^2 - 8\,g_{1,2}\,g_{1,3}{}^3\,g_{2,3} + 2\,g_{1,1}\,g_{1,2}\,g_{1,3}\,g_{2,3}\,g_{3,3} + 4\,g_{2,2}\,g_{1,3}{}^4 + 3\,g_{1,1}\,g_{1,3}{}^2\,g_{2,3}{}^2 - 4\,g_{1,1}\,g_{2,2}\,g_{1,3}{}^2\,g_{3,3}\right)}{4\,g_{1,3}{}^2\left(g_{3,3}\,g_{1,2}{}^2 - 2\,g_{2,3}\,g_{1,2}\,g_{1,3} + g_{2,2}\,g_{1,3}{}^2\right)}, \tag{6.8}$$

where the g's are pupil averages over aberration contributions that can be evaluated analytically and are given in [24]. Solving W_{rms} for a maximum tolerable aberration level W_d gives:

$$\frac{d}{M_{\mathrm{ax}}} = \frac{W_d}{\xi}. \tag{6.9}$$

Figure 6.6 shows d according to this equation for Maréchal's criterion of diffraction limited imaging $W_d = 72\,\mathrm{m}\lambda$. A sample refractive index of $n = 1.5$ and an air objective are assumed. The axial range in object space is provided in units of the diffraction length scale λ/NA^2. This quantity is therefore proportional to the number of independent scan layers that can be acquired, with a constant of proportionality on the order of one. To provide some numerical examples: for a wavelength of $\lambda = 500\,\mathrm{nm}$ an axial range of 67 µm is found for a detection NA = 0.75 and a range of 382 µm is found for a detection NA = 0.50. These values might be sufficient for many use cases where LSM is to be combined with WSI.

6.2.6. Opportunities for computational imaging techniques

In Chapter 3 a deconvolution technique is presented for achieving sectioning in post-processing of multi-focal brightfield images. A simplified 'z-only' algorithm was proposed to achieve a very high processing throughput. In fact, the processing throughput exceeds the acquisition speed by an order of magnitude and can be made even considerably faster by making use of large scale parallelization. The 'time budget' for real time processing therefore has room for a more complex deconvolution technique with potentially a better sectioning of the focal layers and possibly improvement of in-plane resolution as well. A first improvement could be to make the deconvolution algorithm dependent on the lateral length scale of the image features. The challenge then is to find an implementation that is compatible with parallel computation platforms. A potential solution might be to use multi-scale image representations [25]. Decomposition of the image data into a wavelet representation makes it possible to adapt the deconvolution settings to the scale of the wavelet domain [26]. Using the 'z-only' approach on each wavelet domain would keep the advantages of a small deconvolution kernel (low memory usage, and suitability for large scale parallelization).

Further, Chapter 3 contains a system design study of Quantitative Phase Tomography (QPT). It was found that the Point Spread Function (PSF) has a band-pass filtering character in both the axial and lateral direction due to the applied frequency filter and due to the missing cone of the cross-spectral density in the

Fourier domain, respectively. This missing cone is a fundamental limitation resulting from the finite size of the illumination source and the pupil of the imaging system [27]. The image formation might be altered by using more elaborate reconstruction techniques that can, at least partially, recover the missing data. Starting point for addressing this challenge could be Transport of Intensity (TIE) based methods [28, 29], keeping in mind the balance between signal reconstruction and noise amplification, as well as the desire to achieve (near) real-time processing speeds.

Finally, it might be interesting to speculate on the use of image fusion techniques [30] to combine QPT images with widefield fluorescence images. In Fig. 3.8c it was shown that the structural information in the QPT images correlates with the location of fluorescent markers. This opens up the opportunity of combining the advantages of both technologies: QPT images can be obtained at a high speed with a high SNR while widefield fluorescence images provide biological specificity but requires significant integration times. A potential application might be to combine both modalities into a single instrument for high speed fluorescence imaging with a high SNR on a relatively simple system.

References

[1] B. Hulsken. "Method for simultaneous capture of image data at multiple depths of a sample". In: *US Patent No. 9,910,258 B2* (2018) (cit. on p. 134).

[2] B. Hulsken. "Scanning imaging system with a novel imaging sensor with gaps for electronic circuitry". In: *US Patent. No. 10,091,445 B2* (2018) (cit. on pp. 134, 137).

[3] B. Hulsken and S. Stallinga. "Sensor for microscopy". In: *US Patent No. 10,353,190 B2* (2019) (cit. on p. 134).

[4] A. Descloux et al. "Combined multi-plane phase retrieval and super-resolution optical fluctuation imaging for 4D cell microscopy". In: *Nature Photonics* 12.3 (2018), pp. 165–172. DOI: 10.1038/s41566-018-0109-4 (cit. on p. 134).

[5] D. Fu et al. "Quantitative dispersion microscopy". In: *Biomedical Optics Express* 1.2 (2010), p. 347. DOI: 10.1364/BOE.1.000347 (cit. on p. 137).

[6] I. Takayanagi. "CMOS Image Sensors". In: *Image Sensors and Signal Processing for Digital Still Cameras*. Ed. by J. Nakamura. CRC Press, 2017. Chap. 5, pp. 143–178. ISBN: 9781315221083. DOI: 10.1201/9781420026856 (cit. on p. 137).

[7] J. Nakamura. *Image Sensors and Signal Processing for Digital Still Cameras*. Ed. by J. Nakamura. CRC Press, 2017, pp. 1–336. ISBN: 9781315221083. DOI: 10.1201/9781420026856 (cit. on p. 138).

[8] C. Dunsby. "Optically sectioned imaging by oblique plane microscopy". In: *Optics Express* 16.25 (2008), pp. 20306–20316. DOI: 10.1364/OE.16.020306 (cit. on pp. 138, 143, 144).

[9] M. B. Bouchard et al. "Swept confocally-aligned planar excitation (SCAPE) microscopy for high-speed volumetric imaging of behaving organisms". In: *Nature Photonics* 9.2 (2015), pp. 113–119. DOI: 10.1038/nphoton.2014.323 (cit. on pp. 138, 144).

[10] C. B. Müller and J. Enderlein. "Image Scanning Microscopy". In: *Physical Review Letters* 104.19 (2010), p. 198101. DOI: 10.1103/PhysRevLett.104.198101 (cit. on p. 141).

[11] A. G. York et al. "Instant super-resolution imaging in live cells and embryos via analog image processing". In: *Nature Methods* 10.11 (2013), pp. 1122–1126. DOI: 10.1038/nmeth.2687 (cit. on p. 141).

[12] S. Roth et al. "Optical photon reassignment microscopy (OPRA)". In: *Optical Nanoscopy* 2.1 (2013), p. 5. DOI: 10.1186/2192-2853-2-5 (cit. on p. 141).

[13] G. M. De Luca et al. “Re-scan confocal microscopy: scanning twice for better resolution”. In: *Biomedical Optics Express* 4.11 (2013), pp. 2644–2656. DOI: 10.1364/BOE.4.002644 (cit. on p. 141).

[14] R. Heintzmann and C. G. Cremer. “Laterally modulated excitation microscopy: improvement of resolution by using a diffraction grating”. In: *Optical Biopsies and Microscopic Techniques III*. Vol. 3568. January 1999. 1999, pp. 185–196. DOI: 10.1117/12.336833 (cit. on p. 142).

[15] G. E. Cragg and P. T. C. So. “Lateral resolution enhancement with standing evanescent waves”. In: *Optics Letters* 25.1 (2000), pp. 46–48. DOI: 10.1364/OL.25.000046 (cit. on p. 142).

[16] M. G. L. Gustafsson. “Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy”. In: *Journal of Microscopy* 198.2 (2000), pp. 82–87. DOI: 10.1046/j.1365-2818.2000.00710.x (cit. on p. 142).

[17] J. T. Frohn, H. F. Knapp, and A. Stemmer. “True optical resolution beyond the Rayleigh limit achieved by standing wave illumination”. In: *Proceedings of the National Academy of Sciences of the United States of America* 97.13 (2000), pp. 7232–7236. DOI: 10.1073/pnas.130181797 (cit. on p. 142).

[18] R. Heintzmann and T. Huser. “Super-Resolution Structured Illumination Microscopy”. In: *Chemical Reviews* 117.23 (2017), pp. 13890–13908. DOI: 10.1021/acs.chemrev.7b00218 (cit. on pp. 142, 143).

[19] M. G. Gustafsson et al. “Three-Dimensional Resolution Doubling in Wide-Field Fluorescence Microscopy by Structured Illumination”. In: *Biophysical Journal* 94.12 (2008), pp. 4957–4970. DOI: 10.1529/biophysj.107.120345 (cit. on p. 143).

[20] J. Huisken. “Optical Sectioning Deep Inside Live Embryos by Selective Plane Illumination Microscopy”. In: *Science* 305.5686 (2004), pp. 1007–1009. DOI: 10.1126/science.1100035 (cit. on p. 143).

[21] P. J. Verveer et al. “High-resolution three-dimensional imaging of large specimens with light sheet–based microscopy”. In: *Nature Methods* 4.4 (2007), pp. 311–313. DOI: 10.1038/nmeth1017 (cit. on p. 143).

[22] B. C. Chen et al. “Lattice light-sheet microscopy: Imaging molecules to embryos at high spatiotemporal resolution”. In: *Science* 346.6208 (2014). DOI: 10.1126/science.1257998 (cit. on p. 143).

[23] E. J. Botcherby et al. “Aberration-free optical refocusing in high numerical aperture microscopy”. In: *Optics Letters* 32.14 (2007), p. 2007. DOI: 10.1364/OL.32.002007 (cit. on p. 144).

[24] S. Stallinga. “Finite conjugate spherical aberration compensation in high numerical-aperture optical disc readout”. In: *Applied Optics* 44.34 (2005), pp. 7307–7312. DOI: 10.1364/AO.44.007307 (cit. on p. 145).

[25] P. Moulin. “Multiscale Image Decompositions and Wavelets”. In: *The Essential Guide to Image Processing*. Elsevier, 2009, pp. 123–142. DOI: 10.1016/B978-0-12-374457-9.00006-8 (cit. on p. 145).

[26] C. Vonesch and M. Unser. “A Fast Thresholded Landweber Algorithm for Wavelet-Regularized Multidimensional Deconvolution”. In: *IEEE Transactions on Image Processing* 17.4 (2008), pp. 539–549. DOI: 10.1109/TIP.2008.917103 (cit. on p. 145).

[27] N. Streibl. “Three-dimensional imaging by a microscope”. In: *Journal of the Optical Society of America A* 2.2 (1985), p. 121. DOI: 10.1364/JOSAA.2.000121 (cit. on p. 146).

[28] L. Waller et al. “Phase and amplitude imaging from noisy images by Kalman filtering”. In: *Optics Express* 19.3 (2011), pp. 2805–2815. DOI: 10.1364/OE.19.002805 (cit. on p. 146).

[29] E. Bostan et al. “Variational Phase Imaging Using the Transport-of-Intensity Equation”. In: *IEEE Transactions on Image Processing* 25.2 (2016), pp. 807–817. DOI: 10.1109/TIP.2015.2509249 (cit. on p. 146).

[30] S. Li et al. “Pixel-level image fusion: A survey of the state of the art”. In: *Information Fusion* 33 (2017), pp. 100–112. DOI: 10.1016/j.inffus.2016.05.004 (cit. on p. 146).


Garlando

Dankwoord

Op deze plek wil ik iedereen bedanken die heeft bijgedragen aan de totstandkoming van dit proefschrift. Het was een voorrecht te werken tussen vele vriendelijke, enthousiaste, geïnteresseerde en boeiende mensen. Allereerst dank ik mijn promotor Sjoerd Stallinga. Hij heeft mij als leermeester ingewijd in de wetenschap en opgeleid in de optica. Ik ben dankbaar voor de vrijheid ik kreeg in het besteden van mijn tijd en budget, en voor de ruimte om op mijn eigen manier te werken. Sjoerds advies en bemoediging hebben mij veel geholpen. Hij heeft mij veel geleerd over de samenwerking met collega's en onderzoekspartners. De gedetailleerde feedback op mijn schrijven was heel leerzaam en heeft een grote bijdrage geleverd aan de kwaliteit van dit proefschrift. Zijn correcte en zorgzame omgangsvorm blijft een voorbeeld voor mij. Dankbaar ben ik mijn promotor Lucas van Vliet. De jaarlijkse voortgangsgesprekken vormen mijlpalen in het project en hebben bijgedragen aan het tijdig afronden van het promotietraject. Ik bewonder het enthousiasme en de gedrevenheid waarmee Lucas de vakgroep leidde en ik ben dankbaar dat ik deel uit mocht maken van deze groep.

Ik bedank onze onderzoekspartner Philips Digital & Computational Pathology, met name Bas Hulsken, Rik Kneepkens, Jelte Peter Vink, Mathijs Rem, Gert-Jan van den Braak, Aslihan Arslan Carisey en Hans van Wijngaarden. Ik heb genoten van de dagen dat ik welkom was op het kantoor in Best. In het bijzonder dank ik voor de volle week dat ik het lab, de file server, de scanner en Mathijs bezet mocht houden voor mijn onderzoek. Jelte dank ik daarnaast ook voor de introductie bij Philips Research. Ik dank ook de industriële onderzoekspartners ProDrive, met name Leon van de Laar, en Nikon, met name Pierre-Emmanuel Monet en Philippe Baert voor hun deelname aan de gebruikerscommissie en alle aandacht en suggesties die ze daarin hadden voor mijn onderzoek. Ik dank de academische onderzoekspartners van het Erasmus Medisch Centrum (EMC), met name de groep van Arno van Leenders. Aan het begin van mijn promotie gaven Kees Vissers en Peter Riegman mij de unieke gelegenheid om een dag mee te lopen op de Pathologie afdeling om kennis te maken met de beroepspraktijk. Het door mij onverwachte bezoek aan de snijkamer heeft diepe indruk gemaakt. Esther Verhoef en Hein Sleddens hebben geholpen met het prepareren van samples, wat van grote waarde is geweest. Arno van Leenders dank ik voor zijn hulp, advies, en enthousiasme. Ook dank ik de academische partners van het Leids Universitair Medisch Centrum (LUMC), met name de groep van Hans Morreau. Fanny Boyaval heeft een waardevolle bijdrage aan dit proefschrift geleverd door speciaal voor dit onderzoek coupes te prepareren. Ik dank Hans en Fanny voor hun betrokkenheid en de hulp bij het onderzoek in dit proefschrift. Ik dank ook Bob van de Water (LUMC) en Paul van Diest van het Universitair Medisch Centrum Utrecht (UMC) voor hun betrokkenheid in de gebruikerscommissie.

Het onderzoek in dit proefschrift bouwt voort op het werk dat werd gedaan door mijn voorganger Mojtaba Shakeri, waarvoor ik hem bijzonder dankbaar ben.

Het was een voorrecht in het eerste halfjaar samen te kunnen werken aan de opstelling, zodat ik een snelle start kreeg in het experimentele werk. Ik heb veel geleerd van zijn zorgvuldige en creatieve aanpak. Veel valkuilen heb ik kunnen vermijden doordat hij de weg al had verkend.

De Imaging Physics afdeling, en met name de Quantitative Imaging groep is de afgelopen jaren als een thuis voor mij geweest. Ik dank de stafleden van de groep voor de vriendschappelijke en collegiale omgang. Het was een voorrecht de promovendi te mogen vertegenwoordigen bij de staff meetings. Ik dank Bernd Rieger voor zijn ongezouten en soms ongevraagde advies, het is mij tot voordeel geweest. Dank ook voor het onderzoek dat ik kon doen naar digitaal toetsen bij Systemen en Signalen, ook al bleek het geen succes. Ook Jeroen Kalkman mocht ik assisteren, bij de colleges Computational Physics en Imaging Systems. Ik dank Ted Young voor het delen van zijn kennis en decennialange ervaring in de microscopie en beeldvorming. Piet Verbeek dank ik voor de interessante gesprekken, onder meer over kleurtheorie. Ook Frans Vos, Jaap Caro en Dirk Poot dank ik voor hun betrokkenheid binnen de vakgroep.

De groep zou zichzelf niet zijn zonder de onmisbare ondersteuning van de support staff. Annelies van Beek en Angela van der Sande vormen een ijzersterk duo en zorgen dat alles in de groep als op rolletjes loopt. Hun kantoor stond altijd voor mij open voor het vinden van de hazenpaden in de universitaire organisatie, maar bovenal voor een persoonlijk gesprek, of gewoon een lekker spekje. Ronald Ligteringen dank ik als een bijzonder betrokken systeembeheerder die altijd open stond voor mijn vreemde computervragen. Zijn grote hart voor mensen en oprechte, opgetogen en openhartige levenshouding waardeer ik zeer. Ron Hoogerheide is een artiest op de draaibank. Ik dank hem voor de vele klussen die hij vaak met voorrang voor mij verrichtte, hij heeft daarmee een grote bijdrage aan mijn experimentele opstelling geleverd. Ook dank ik de heren van DEMO, met name Hans van der Does en Jack Maat, waar ik altijd kon binnenlopen met elektronische vraagstukken of voor het 'lenen' van componentjes. Ook bij Henry den Bok kon ik altijd aankloppen voor ontbrekende componenten of technisch advies. Ik dank Raymond Schouten voor het delen van zijn bijzonder specialistische elektronica kennis, met name met betrekking het zeer snel schakelen van LEDs. Nicolette Meerstadt, Joyce Alexander-Buis en Liesbeth Secker dank ik voor hun goede zorg voor de afdeling en de vrolijke sfeer.

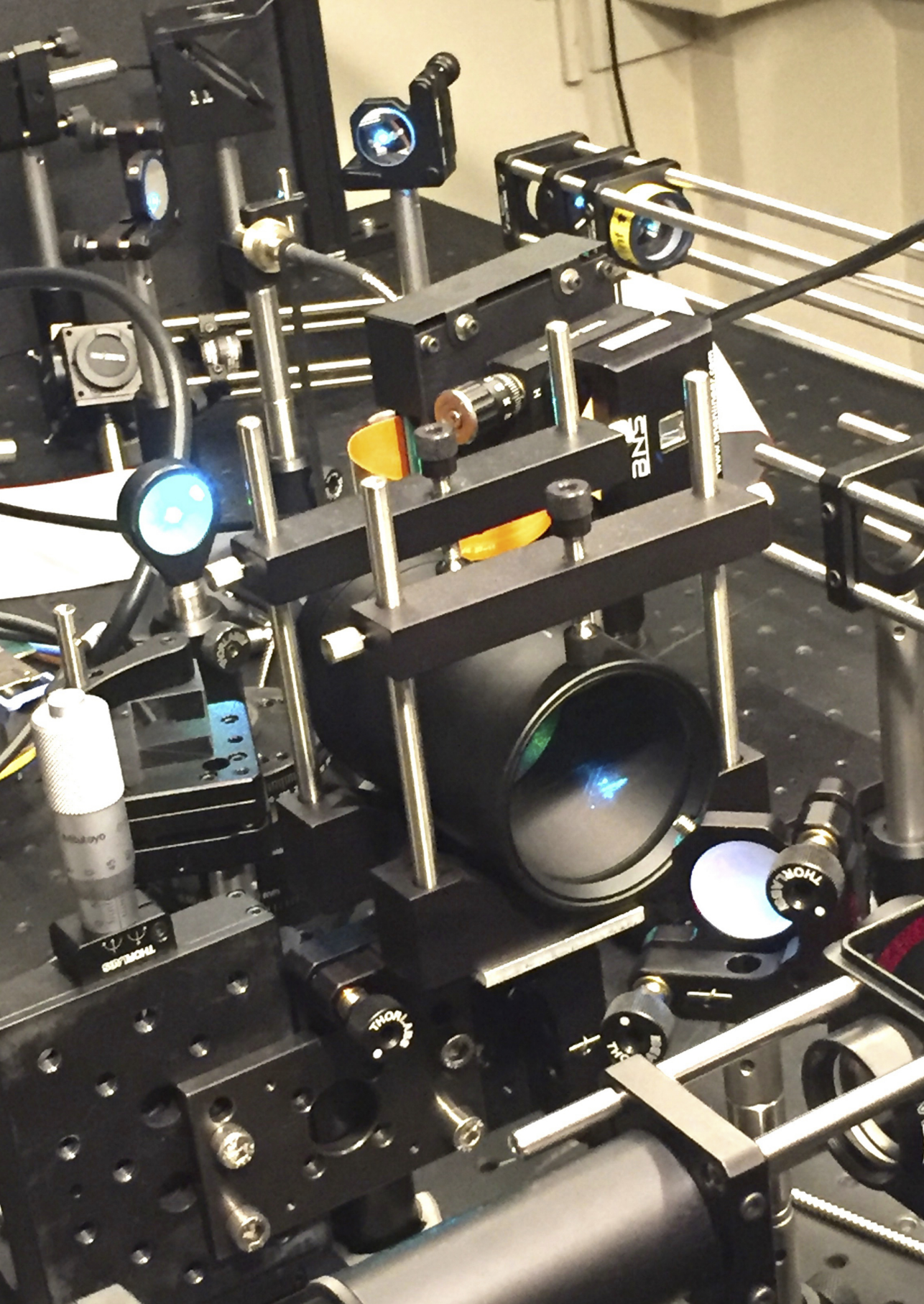
Ik ben dankbaar voor de samenwerking met de promovendi van Quantitative Imaging (QI). Ik heb veel geleerd van Jelle van der Horst, die bij veel optomechanische ontwerpen heeft geadviseerd. Zijn opgewekte levenshouding en collegialiteit hebben voor mij veel bijdragen aan de goede sfeer in QI. Ook Marijn Siemons dank ik voor zijn betrokkenheid bij, en bijdrage aan het experimentele werk. Gyllion Loozen dank ik voor de kameraadschap en de vele leerzame discussies. Ik bewonder de leergierigheid waarmee hij zich weet te storten op alles wat hem interesseert, of dat nu Chinees, filosofie of optica is. Jelle Storteboom dank ik voor het delen van zijn laser expertise. Ook de experimentalisten Christiaan Hulleman, Robert Moerland, Taylor Hinsdale, Jeroen Heldens, Nadia Chakrova, Boling Ouyang, Jos van Rooij en Arnica Karuna dank ik voor hun advies. Een aantal jaren heb ik doorgebracht met Tom Callewaert als buurman, zowel in het lab als het kantoor, en heb genoten van de genuanceerde gesprekken en sarcastische grappen. Ik vind

het heel leuk dat Jos de Wit het laatste jaar het kantoor met mij deelde. Wij leerden elkaar al kennen in de collegebanken en hebben tijdens de studie vele uren samen sommetjes gemaakt. In de afronding van mijn promotie waren de discussies met hem heel motiverend. Verreweg de langste tijd deelde ik het kantoor met Yan Guo, wat mij een voorrecht is geweest. Ik bewonder haar vermogen tot het stellen van verrassende vragen over de dingen die ik als alledaags aanneem. Ook heb ik jarenlang de kamergenoot mogen zijn van Anna Katharina Trull. Ik dank haar voor alle gedeelde ervaringen, en biedt bij dezen excuses aan dat ik mijn kamerplant naar haar heb vernoemd. Ook dank ik kamergenoot Emile Broeders voor de goede tijd en de wiskunde die hij mij heeft uitgelegd. Ook groepsgenoten Pierre Ambrosini, Lena Filatova, Rasmus Thorsen, Hamidreza Heydarian, Willem van Valenberg, Juan Pedro Vigueras Guillén, Tian Zhang en Martijn Nagtegaal dank ik voor de goede tijd bij QI. Ik heb veel plezier beleefd en geleerd van de samenwerking en begeleiding van studenten, waarvan ik Joost de Jong, Marten Kruidhof, Derk van Grootheest, Sam Pulles, Sten Goes, en Adriaan Visser met name wil noemen.

Ik ben dankbaar voor de jarenlange vorming die voorafging aan het schrijven van dit proefschrift. Daarom benut ik deze gelegenheid graag om te danken voor het basis- en middelbaar onderwijs dat ik genoot. Ik wil met name basisschoolleraar Rob van Asperen danken voor zijn enthousiasme en ongeëvenaarde schrijflessen. Hij was de eerste die mij ooit vertelde over promoveren, met de uitleg 'dan moet je een boek schrijven over iets waar nog niemand eerder iets over geschreven heeft.' Ook dank ik middelbare school docent in de natuurkunde Dick Vierbergen die mij heeft geïntroduceerd in dit schitterde vakgebied.

Bijzonder dankbaar ben ik mijn ouders voor het goede thuis, de liefdevolle opvoeding en de blijvende trots en aanmoediging. Zij hebben mijn enthousiasme voor wetenschap en techniek altijd aangemoedigd en gedeeld. Ik dank mijn schoonouders voor alle betrokkenheid en de dagen dat ik bij hen heb mogen werken aan dit proefschrift. Ook wil ik hier Hilde bedanken, met haar op mijn knie had ik één van de betere ideeën voor dit proefschrift. Ik dank Gerdien dat ik met haar alle plezier van dit promotietraject heb kunnen delen en dat zij altijd heel betrokken is geweest op het leven op de universiteit. Dat ik altijd mag rekenen op haar vertrouwen en aanmoediging is voor mij volstrekt onmisbaar. Tenslotte gaat al mijn dank uit naar God, de vader, schepper van hemel en aarde, die alle zichtbare en onzichtbare dingen heeft gemaakt.

Leon van der Graaff,
oktober 2019



Curriculum Vitæ

Leon van der Graaff was born on June 6, 1988 in Rotterdam, The Netherlands. He obtained his high school diploma in 2006 at Melanchthon Rotterdam. Subsequently he subscribed for the bachelor Applied Physics at the Delft University of Technology. During his bachelor he followed the education minor, including internship as a teacher at the Ashram College in Alphen aan den Rijn, resulting in a second grade teaching degree in physics. Also, he spend one year in a full time position as Treasurer at the student union C.S.R. Delft. The bachelor was concluded with a thesis on the design of a fluorescence illumination unit for a whole slide scanning system. A BSc degree was obtained in 2013. His studies were continued with a master Applied Physics at same university. During his master he followed the education track and did an internship as a teacher at Melanchthon Bergschenhoek and obtained a first grade teaching degree in physics. In his master project experimental research was performed on new acquisition protocols for medical ultrasound for added contrast between blood and tissue. A MSc title was obtained in 2015. During his studies, he owned a one-man business in the development, design and technical realization of websites for SME's. September 2015 he started his PhD in the department of Imaging Physics at the Delft University of Technology, resulting in this thesis. Starting January 2020 he works as a Scientist Optics at Philips Research, Eindhoven.

Fig. 1. Chan
changing the
the free work

disc, finit
pensate
around
this me
known
betwe
the n
conju
of th
with
me
fec
co
co
r
t

List of Publications

paper — L. van der Graaff, G.J.L.H. van Leenders, F. Boyaval, S. Stallinga, "Computational imaging modalities for multi-focal whole slide imaging systems," *Applied Optics* **59**, 20 (2020). DOI: 10.1364/AO.394290

paper — L. van der Graaff, G.J.L.H. van Leenders, F. Boyaval, S. Stallinga, "Multi-line fluorescence scanning microscope for multi-focal imaging with unlimited field of view," *Biomedical Optics Express* **10**, 12 (2019). DOI: 10.1364/BOE.10.006313

patent — L. van der Graaff, S. Stallinga, "Multi-line confocal scanner," International Patent PCT/EP2019/074698 (filed 16 Sep 2019).

presentation — L. van der Graaff, S. Stallinga, "YES WE SCAN," OSA Applied Industrial Optics, Biophantastic! W2A.5 (15 April 2019)

presentation — L. van der Graaff, S. Stallinga, "Confocal Multi-Line Scanning Microscope for Efficient 3D Fluorescence Imaging," Focus on Microscopy, MO-MO-PAR-D High Speed/Content/Throughput (15 April 2019)

presentation — L. van der Graaff, S. Stallinga, "3D fluorescence whole slide imaging using confocal multi-line scanning," QBI, Digital Pathology session, #93 (11 Jan 2019)

presentation — L. van der Graaff, S. Stallinga, "3D fluorescence whole slide imaging using confocal multi-line scanning," EOSAM, TOM 4 Bio-Medical Optics (11 Oct 2018)

in proceedings — L. van der Graaff, G.J.L.H. van Leenders, F. Boyaval, S. Stallinga, "Fluorescence imaging for whole slide scanning using LED-based color sequential illumination," Proc. SPIE 10679, Optics, Photonics, and Digital Technologies for Imaging Applications V, 106790D (24 May 2018) DOI: 10.1117/12.2306776

poster — L. van der Graaff, S. Stallinga, "Fluorescence Whole Slide Imaging System Based on Color-Sequential Led-Based Line Scanning," Conference poster, Focus on Microscopy, P1-F Fluorescence I (10 April 2017)

Propositions

accompanying the dissertation

Multi-modal Whole Slide Imaging

by

Leon van der Graaff

1. 3D confocal Whole Slide Imaging is most efficiently implemented by a multi-focal line scanning approach (this thesis).
2. Multi-line scanning with a tilted image plane does not require a dedicated sensor (this thesis).
3. Quantitative Phase Tomography (QPT) is an ideal candidate for label-free Whole Slide Imaging for its low demands on the hardware and its simple algorithm (this thesis).
4. Deep insight in something is not only shown by answering complicated questions, but also by explaining it simple. The layman's talk should therefore be an integral part of a PhD defence ceremony.
5. Direct instruction is a valuable teaching method that cannot be replaced by self study.
6. 'The mantra "Little changes can make a big difference" is bunkum, when applied to climate change and power. Every BIG helps' [1].
7. 'When a measure becomes a target, it ceases to be a good measure' (Goodhart's law) [2].
8. A doctor should be a subscriber of a quality newspaper to be able to form a judgement society can rely on.
9. Online targeted ads are a threat to the open society.
10. *Gezelligheid* requires shared responsibility.

These propositions are regarded as opposable and defendable, and have been approved as such by the promotors
prof. dr. S. Stallinga and prof. dr. ir. L. J. van Vliet.

[1] D. J. C. MacKay. *Sustainable Energy — without the hot air*. Uit, 2008.
[2] M. Strathern. "'Improving ratings': audit in the British University system". In: *European Review* 5.03 (1997), pp. 305–321.

Stellingen

behorende bij het proefschrift

Multi-modal Whole Slide Imaging

door

Leon van der Graaff

1. 3D confocale microscopie van hele voorwerpglaasjes kan het meest efficient worden gedaan met multifocaal lijnscannen (dit proefschrift).
2. Multi-lijn scannen met een gekanteld beeldvlak heeft geen speciaal daarvoor ontwikkelde beeldsensor nodig (dit proefschrift).
3. *Quantitative Phase Tomography* (QPT) is een ideale kandidaat voor het inscannnen van grote stukken ongelabeled weefsel omdat het weinig eisen stelt aan het instrument en omdat het algoritme eenvoudig is.
4. Diep inzicht blijkt niet alleen uit het beantwoorden van moeilijke vragen, maar ook uit de vaardigheid iets eenvoudig uit te leggen. Het lekenpraatje moet daarom integraal onderdeel van de promotieplechtigheid worden.
5. Directe instructie is een waardevolle onderwijsvorm die niet kan worden vervangen door zelfstudie.
6. 'Het mantra "Alle kleine beetjes helpen" is gebazel als het gaat over klimaatverandering en energie. Alle GROTE beetjes helpen' [1].
7. 'Wanneer een maat een doel wordt, is het geen goede maat meer' (Goodhart's law) [2].
8. Een doctor heeft een abonnement op een kwaliteitskrant nodig om een oordeel te kunnen vormen waarop de maatschappij kan vertrouwen.
9. Gepersonaliseerde online advertenties zijn een bedreiging voor de open samenleving.
10. Gezelligheid vereist gedeelde verantwoordelijkheid.

Deze stellingen worden opponeerbaar en verdedigbaar geacht en zijn als zodanig goedgekeurd door de promotoren prof. dr. S. Stallinga en prof. dr. ir. L. J. van Vliet.

[1] D. J. C. MacKay. *Sustainable Energy — without the hot air*. Uit, 2008.
[2] M. Strathern. "'Improving ratings': audit in the British University system". In: *European Review* 5.03 (1997), p. 305–321.