ARRAY CONFOCAL MICROSCOPY

by

Shaun Pacheco

A Dissertation Submitted to the Faculty of the

COLLEGE OF OPTICAL SCIENCES

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2017

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation prepared by Shaun Pacheco, titled Array Confocal Microscopy and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

_____ Date: 12/02/2016
Rongguang Liang

_____ Date: 12/02/2016
Tom Milster

_____ Date: 12/02/2016
Leilei Peng

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

_____ Date: 12/02/2016
Dissertation Director: Rongguang Liang

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of the requirements for an advanced degree at the University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that an accurate acknowledgement of the source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: Shaun Pacheco

ACKNOWLEDGEMENTS

First and foremost, I would like to express my sincere gratitude to my advisor Dr. Rongguang Liang for his support throughout my entire Ph.D. career. I am deeply thankful to him for giving me the tools and the support for letting me pursue a wide array of projects throughout my Ph.D. career. His support and guidance has helped shape me into the researcher that I am today.

I am extremely thankful to the other members of my committee Dr. Tom Milster and Dr. Leilei Peng for their useful comments and discussions. Dr. Tom Milster's guidance was instrumental in my research in diffractive optics. In addition to several useful discussions about biomedical imaging, Dr. Leilei Peng taught me how to properly program in LabView, which was instrumental in building array confocal microscopes.

In particular I would like to thank Lee Johnson and Melissa Zaverton for fabricating beam-splitting gratings that are described in this dissertation. I would also like to thank Chengliang Wang, Urs Utzinger, Minkkhoi Nguyen, Monica Chawla, and Carol Barnes for their help in developing an array confocal microscope for whole brain imaging.

I would like to thank everyone I've collaborated with throughout my Ph.D, in particular I would like to thank: Zhenyue Chen, Nitigya Kathuria, Basel Salahieh, Donglin Ma, Xiaoyin Zhu, Ross Uthoff, Dmitry Reshidko, Weichuan Gao, Jonathan Brand, Eric Herman, Chih Yu Huang, Nan Zhu, Zexin Feng, Rengmao Wu, and Tingkui Mu for their useful discussions and collaborations. I am grateful to all the members of the Applied Optics Lab that I have worked with over the years. It has been a privilege collaborating on various projects and having the opportunity to learn from each other.

Special thanks to the professors and the teaching assistants from all the courses I've taken at University of Arizona. My Ph.D. would not be possible without learning from their expertise. I would like to thank the faculty and staff at University of Arizona, without whom the university would not function properly.

I am grateful to all my friends for their friendship and the memories we've made during graduate school.

Finally, I would like to give my heartfelt thanks to my family, who has consistently provided me with love, encouragement, and support my entire life.

This research was supported by several agencies: Biomedical Imaging and Spectroscopy Fellowship (BMIS), ARCS Fellowship, National Institutes of Health, McKnight Brain Research Foundation, Office of Research and Discovery at University of Arizona, BIO5 Institute, State of Arizona, and ADHS.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Confocal microscopes utilize point illumination and pinhole detection to reject out-of-focus light. Because of the point illumination and detection pinhole, confocal microscopes typically utilize point scanning for imaging, which limits the overall acquisition speed. Due to the excellent optical sectioning capabilities of confocal microscopes, they are excellent tools for the study of three-dimensional objects at the microscopic scale. Fluorescence confocal microscopy is especially useful in biomedical imaging due to its high sensitivity and specificity. However, all designs for confocal microscopes must balance tradeoffs between the numerical aperture (NA), field of view (FOV), acquisition speed, and cost during the design process.

In this dissertation, two different designs for an array confocal microscope are proposed to significantly increase the acquisition speed of confocal microscopes. An array confocal microscope scans an array of beams in the object plane to parallelize the confocal microscope to significantly reduce the acquisition time. If N beams are used in the array confocal microscope, the acquisition time is reduced by a factor of N. The first design scans an array of miniature objectives over the object plane to overcome the trade-off between FOV and NA. The array of objectives is laterally translated and each objective scans a small portion of the total FOV. Therefore, the number of objectives used in the array limits the FOV, and the FOV is increased without sacrificing NA. The second design utilizes a single objective with a high NA, large FOV, and large working distance designed specifically for whole brain imaging. This array confocal microscope is designed to speed up the acquisition time required for whole brain imaging. Utilizing an

objective with a large FOV and scanning using multiple beams in the array significantly reduces the time required to image large three-dimensional volumes.

Both array confocal microscope designs use beam-splitting gratings to efficiently split one laser beam into a number of equal energy outgoing beams, so this dissertation explores design methods and analyses of beam-splitting gratings to fabrication errors. In this dissertation, an optimization method to design single layer beam-splitting gratings with reduced sensitivity to fabrication errors is proposed. Beam-spitting gratings are typically only designed for a single wavelength, so achromatic beam-splitting grating doublets are also analyzed for possible use in array confocal microscopes with multiple excitation wavelengths. An analysis of the lateral shift between grating layers in the achromatic grating doublet proves grating profiles with constant first spatial derivatives are significantly less sensitive than continuous phase profiles. These achromatic grating doublets have designed performance at two wavelengths, but the diffraction angles at the two wavelengths differ. To overcome that limitation, scale-invariant achromatic gratings are designed, which not only provide designed performance at two wavelengths, but also equal diffraction angles at two wavelengths.

1. INTRODUCTION TO CONFOCAL MICROSCOPY

1.1 Introduction

The goal of a microscope is to produce a magnified image of a microscopic sample without degrading the image quality. Microscopes have become an essential tool for many biomedical applications. They are used in investigating biological processes, diagnosing diseases, and quantitatively measuring biological processes *in vitro* and *in vivo*. Several imaging techniques utilizing microscopes are prevalent in biomedical imaging such as bright field microscopy, dark field microscopy, phase contrast microscopy, fluorescence imaging, light sheet microscopy, multiphoton microscopy, deconvolution microscopy and confocal microscopy. Due to the wave nature of light, the resolution of a microscope using these techniques is limited by diffraction. Recently, several superresolution techniques, such as STED [1-4], STORM [5-8], PALM [9,10], linear SIM [11-14] and nonlinear SIM [15,16] have been developed to probe features smaller than the diffraction limit. Confocal microscopes are of particular interest due to their excellent optical sectioning capabilities. The optical sectioning capabilities allows for high-resolution images inside scattering media, which is an extremely useful tool for several biomedical imaging applications.

This chapter introduces the concept of diffraction-limited microscopy by describing a bright field microscope. Then a confocal microscope is described by highlighting the improvements over a bright field microscope. Finally, the advantages of array confocal microscopy are described.

1.2 Bright Field Microscopy

1.2.1 Optical Layout

A typical optical layout for a microscope with an infinity-corrected objective is shown in Fig. 1.1. The object plane is at the front focal plane of the objective, and the output for an infinity-corrected objective is a collimated beam for every object point. A tube lens is used to form an intermediate image, which can be directly imaged onto an electronic sensor or observed by the human eye through the eyepiece. The exit pupil of the objective lens is typically set at the rear focal plane to make the objective object-space telecentric. In a telecentric system, the chief rays are parallel to the optical axis and the system magnification is constant even if the object is displaced from the focal plane. Microscope objectives are well corrected for aberrations, and thus produce diffraction-limited imaging.

Two important properties of a microscope are the numerical aperture (NA) and magnification. The NA of the microscope objective is defined as

$$\mathrm{NA} = n \sin \alpha \tag{1.1}$$

where n is the refraction index of the medium between the front lens of the objective and the object, α is half acceptance angle of the objective. The magnification of the objective is defined as

$$M_{obj} = \frac{f_{tube}}{f_{obj}}, \tag{1.2}$$

where f_{tube} is the focal length of the tube lens and f_{obj} is the focal length of the objective. The total magnification of the microscope with an eyepiece is the product of the magnification of the objective and the magnification of the eyepiece

$$M_{microscope} = M_{obj} M_{eye}. \tag{1.3}$$

The magnification of the eyepiece is approximately

$$M_{eye} \approx \frac{250\ \text{mm}}{f_{eye}}, \tag{1.4}$$

where f_{eye} is the focal length of the eyepiece.

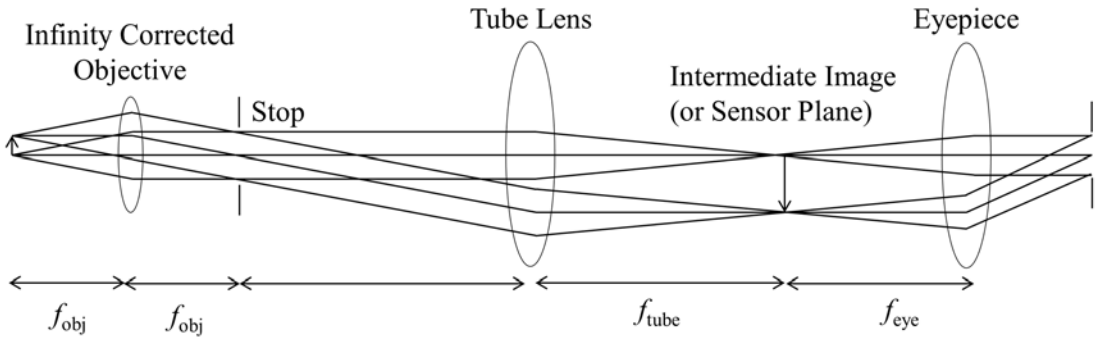

Fig. 1.1 A microscope design for an infinity corrected objective.

1.2.2 Resolution

The complex exit pupil of the objective is defined as

$$P(x,y) = A(x,y)\exp\left(i\tfrac{2\pi}{\lambda}W(x,y)\right), \tag{1.5}$$

where A is the amplitude function and W is the aberration function of the objective at the exit pupil for wavelength λ. Diffraction from the exit pupil to the plane of focus yields the impulse response

$$h(x_i, y_i) = \iint_{-\infty}^{\infty} P(x,y)\exp\left(-i2\pi\left(\frac{x}{\lambda d_i}x_i + \frac{y}{\lambda d_i}y_i\right)\right)dxdy, \tag{1.6}$$

where d_i is the distance from the exit pupil to the image plane, and x_i and y_i are the spatial coordinates at the image plane. The impulse response is the Fourier transform of the complex pupil function.

For an aberration free objective with a circular aperture, the pupil function $P(x,y)$ is unity, the point spread function (PSF) is:

$$|h(r_i)|^2 = \left(\frac{\pi R^2}{\lambda z}\right)^2 \left(2\frac{\mathrm{J}_1\left(\frac{2\pi NAr_i}{\lambda}\right)}{\frac{2\pi NAr_i}{\lambda}}\right)^2, \tag{1.7}$$

where r_i is the radial coordinate in image space, R is the radius of exit pupil, and J_1 is the Bessel function of the first kind. The normalized irradiance distribution of the diffraction-limited PSF is shown in Fig. 1.2. The first zero of the Airy pattern is at a radial distance

$$r_i = \frac{0.61\lambda}{NA}. \tag{1.8}$$

Rayleigh's criterion is often used as a measure of the resolution in microscopy. Rayleigh's criterion states that two incoherent point sources are just resolvable when the center of one falls exactly on the null of the second, as shown in Fig. 1.3. This corresponds to a dip between the two peaks of approximately 74% of the maximum. All sources closer than Rayleigh's criterion cannot be resolved, and are assumed to come from the same point source.

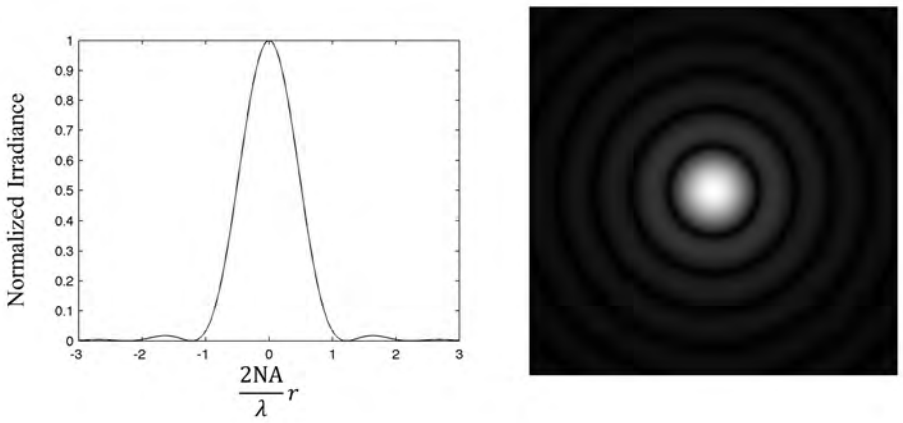

Fig. 1.2 The lateral and 2D profile of the diffraction-limited point spread function.

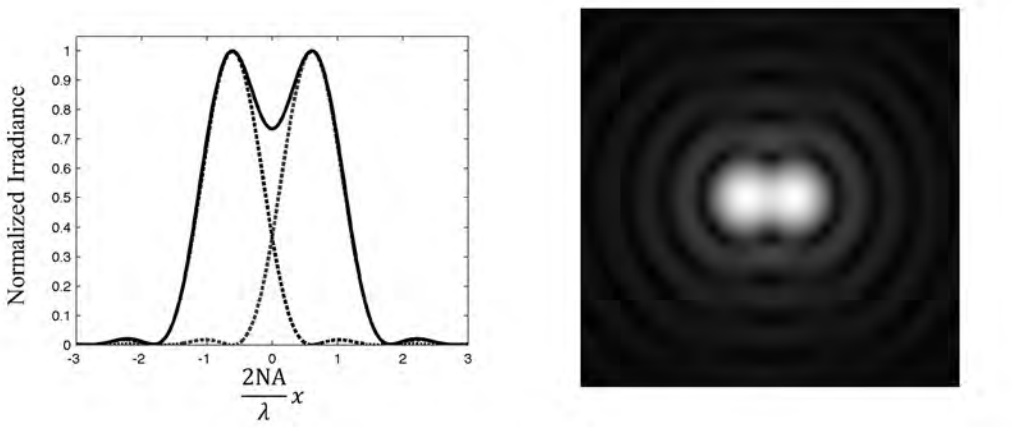

Fig. 1.3 The lateral and 2D profile of Rayleigh's criterion.

For incoherent illumination, the intensity of the image is given by

$$I_i(x,y) = I_g(x,y) * \left|h(x,y)\right|^2, \tag{1.9}$$

where * is the convolution operator, $|h|^2$ is the incoherent PSF of the optical system and I_g is the image as predicted by geometrical optics. The resulting image is a blurring of the geometrical optics image by the point spread function.

Since an optical system can be modeled as a linear shift invariant system, the imaging equation can be written in the frequency domain. The imaging equation in frequency space is

$$G_i(\xi,\eta) = \mathrm{OTF}(\xi,\eta) G_g(\xi,\eta), \tag{1.10}$$

where ξ and η are the spatial frequency coordinates, G_i is the normalized Fourier transform of the image, G_g is the normalized Fourier transform of the image predicted by geometrical optics, and OTF is the optical transfer function of the imaging system. OTF is the normalized Fourier transform of the PSF:

$$\mathrm{OTF}(\xi,\eta) = \frac{\iint \left|h(x,y)\right|^2 \exp\left(-i2\pi(\xi x + \eta y)\right) dxdy}{\iint \left|h(x,y)\right|^2 dxdy}. \tag{1.11}$$

Some important properties of the OTF are:

1. OTF(0,0) = 1,
2. OTF(ξ, η) = OTF*(-ξ,-η),
3. |OTF(ξ, η)| ≤ |OTF(0,0)| = 1.

The OTF for a diffraction-limited system is shown in Fig. 1.4, it has a cutoff frequency at 2NA/λ. Objects with a higher spatial frequency cannot be resolved, unless superresolution techniques are used.

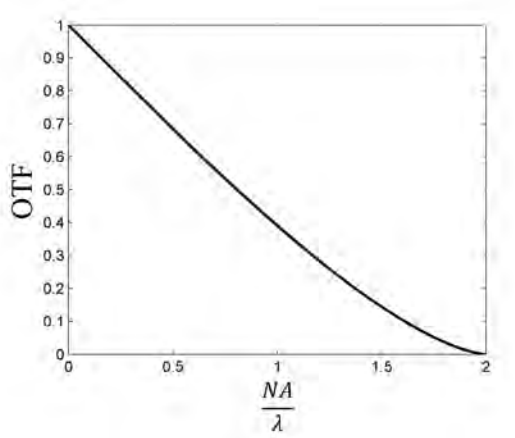

Fig. 1.4 Diffraction-limited OTF.

1.2.3 Fluorescence Microscopy

Instead of detecting the reflectance, transmission, or absorption of light, fluorescence microscopy utilizes fluorescence to study samples under a microscope. Fluorescence is the process by which an atom or a molecule emits light at a longer wavelength than the illumination wavelength. A sample can be stained by a fluorescent dye to investigate specific structures of interest in a biological sample. Furthermore, a sample can be stained with several different fluorescent dyes that fluoresce at different wavelengths. Due to the different fluorescent wavelengths, multiple structures of interest can be studied with high specificity, as long as the emitted light can be effectively differentiated upon detection. The fluorescence signal is typically orders of magnitude smaller than the excitation; therefore, optical filters are used to separate the excitation and fluorescence

signal in a microscope. Confocal microscopes can utilize either reflectance or fluorescence microscopy to image a 3D volume of interest.

1.3 Confocal Microscopy

One disadvantage of bright field microscopy is light from the out-of-focus region of the image is overlapped with the focused region at the image plane. This results in poor contrast in the final image. Confocal microscopy was developed by Minsky in 1957 [17]. Confocal microscopes place a pinhole in front of a detector to reject the out-of-focus light, which leads to optical sectioning and improved contrast over bright field microscopy. Due to the optical sectioning, a confocal microscope is used to image a volume noninvasively. This section discusses the principle of confocal microscopy and its improvement over conventional bright field microscopes. The components used in a confocal microscope and some modifications to the design are also discussed.

1.3.1 Principle

The principle of a confocal microscope is shown in Fig. 1.5. Light travels through an illumination pinhole and is focused onto a sample on the focal plane of the objective. The reflected and/or fluorescent light at the focal plane returns through the objective and is focused at the image plane, where a detection pinhole is placed, as shown in Fig. 1.5(a). The pinhole has the function of preventing the majority of the out-of-focus light from reaching the detector. The returning light from the focal point passes through the small detection pinhole, where it is finally detected by the detector. As shown in Fig. 1.5(b), the light from the axial plane above and below the focal plane focuses either above or below the detection pinhole and the majority of the out-of-focus light is blocked by the pinhole. Additionally, the pinhole blocks the light from points on the focal plane that are laterally

shifted from the illumination spot as shown in Fig. 1.5(c). Since the out-of-focus light is rejected, confocal microscopes have optical sectioning capabilities and have increased contrast over bright field microscopes. Due to the pinhole, confocal microscopes only image a single point onto a detector at a time. Therefore, a scanning system is required to scan the focused beam across the object of interest.

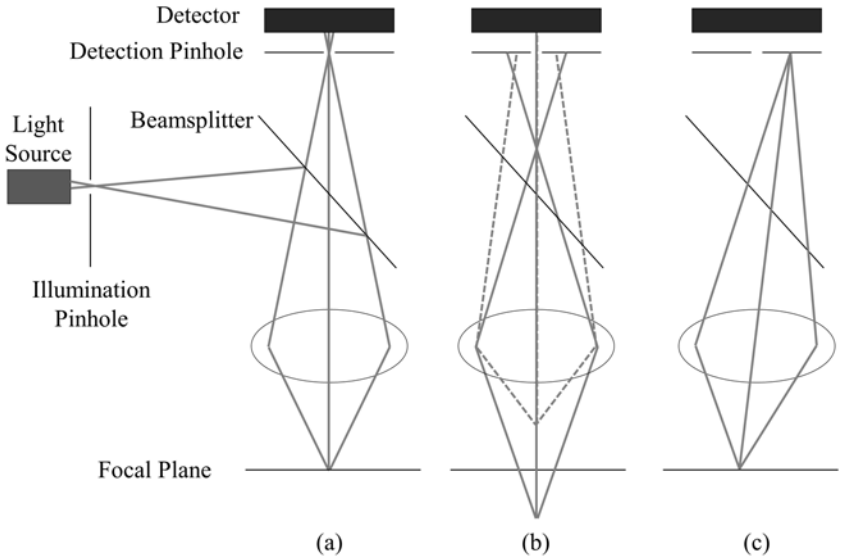

Fig. 1.5 Concept for confocal microscope. (a) Focused beam is transmitted through the detection pinhole. (b) The majority of the out-of-focus light is blocked by the detection pinhole. (c) Light that is laterally shifted from the focused beam is also blocked by the pinhole.

The PSF behind the detection pinhole for a confocal microscope is the product of the PSFs of the illumination path and the detection path:

$$\mathrm{PSF}_{tot}(r) = \mathrm{PSF}_{ill}(r)\mathrm{PSF}_{det}(r), \tag{1.12}$$

where PSF_{ill} and PSF_{det} are the PSFs for the illumination and detection path, respectively. Consider a diffraction-limited confocal microscope, the total PSF is proportional to

$$\mathrm{PSF}_{tot}(r) \propto \mathrm{somb}^2\left(\frac{2\mathrm{NA}_{ill}}{\lambda_{ill}}r\right)\mathrm{somb}^2\left(\frac{2\mathrm{NA}_{det}}{\lambda_{det}}r\right), \tag{1.13}$$

where NA_{ill} and λ_{ill} is the NA and wavelength for the illumination path, and NA_{det} and λ_{det} is the NA and wavelength for the detection path. For a reflectance confocal system, the

illumination and detection path utilize the same objective and wavelength, so the PSF of the confocal system is

$$\mathrm{PSF}_{tot}(r) \propto \mathrm{somb}^4\left(\tfrac{2\mathrm{NA}}{\lambda} r\right). \tag{1.14}$$

Therefore, the PSF of a confocal microscope is the square of the Airy disk. A comparison of the lateral profile of the PSF for a bright field and confocal microscope is shown in Fig. 1.6(a). While the first zeros of the PSFs are at the same location, Fig. 1.6 shows the full width at half maximum (FWHM) for a confocal microscope is smaller than conventional microscopes. The resolution for a confocal microscope according to Rayleigh's criterion is [18]

$$\Delta r_{confocal} = \frac{0.51\bar{\lambda}}{NA}, \tag{1.15}$$

where

$$\bar{\lambda} = \sqrt{2}\,\frac{\lambda_{exc}\lambda_{em}}{\sqrt{\lambda_{exc}^2 + \lambda_{em}^2}}. \tag{1.16}$$

Furthermore, the axial PSF of a confocal microscope is proportional to [19]

$$\mathrm{PSF}(z) \propto \mathrm{sinc}^2\left(\tfrac{NA_{ill}^2}{2\lambda_{ill}} z\right)\mathrm{sinc}^2\left(\tfrac{NA_{det}^2}{2\lambda_{det}} z\right). \tag{1.17}$$

Figure 1.6(b) compares the axial PSF of the confocal microscope to conventional microscopes. The axial resolution from Rayleigh's criterion is [18]

$$\Delta z_{confocal} = \frac{0.88\bar{\lambda}}{\left(n - \sqrt{n^2 - \mathrm{NA}^2}\right)}. \tag{1.18}$$

Figure 1.7 shows *x*-*z* slices of the lateral and axial resolution. Note the 3D PSF for the confocal microscope is narrower than the PSF for conventional microscopes in all 3 dimensions; therefore the resolution is improved laterally and axially. Since the pinhole

blocks the out-of-focus light, confocal microscopes have optical sectioning capabilities also. The optical sectioning allows the confocal microscope to capture 3D images of an object of interest. This discussion assumes the pinhole is infinitely small, when the pinhole is increased; the lateral and axial resolution decreases [20].

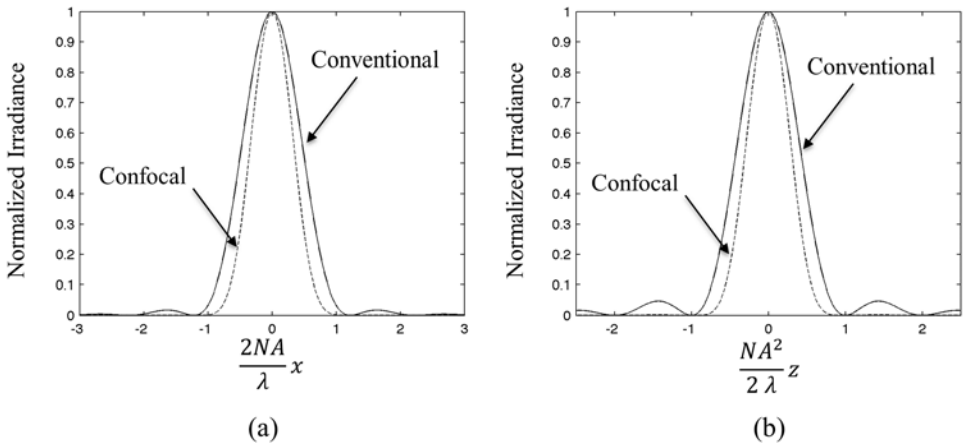

Fig. 1.6 A comparison of the (a) lateral PSF and (b) axial PSF for conventional microscopy with confocal microscopy.

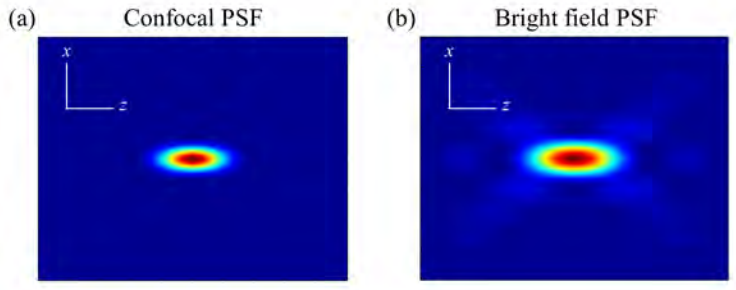

Fig. 1.7 The PSF in the *x*-*z* plane for (a) confocal microscopy and (b) bright field microscopy.

1.3.2 Components

Although the system design for a confocal microscope varies depending on the specific application, a representative example layout for a confocal microscope is shown in Fig. 1.8. Confocal microscopes consist of a light source, an illumination pinhole, illumination optics, a beamsplitter, a scanning mechanism, one or two relay systems, a microscope objective, detection optics, a detection pinhole and a detector.

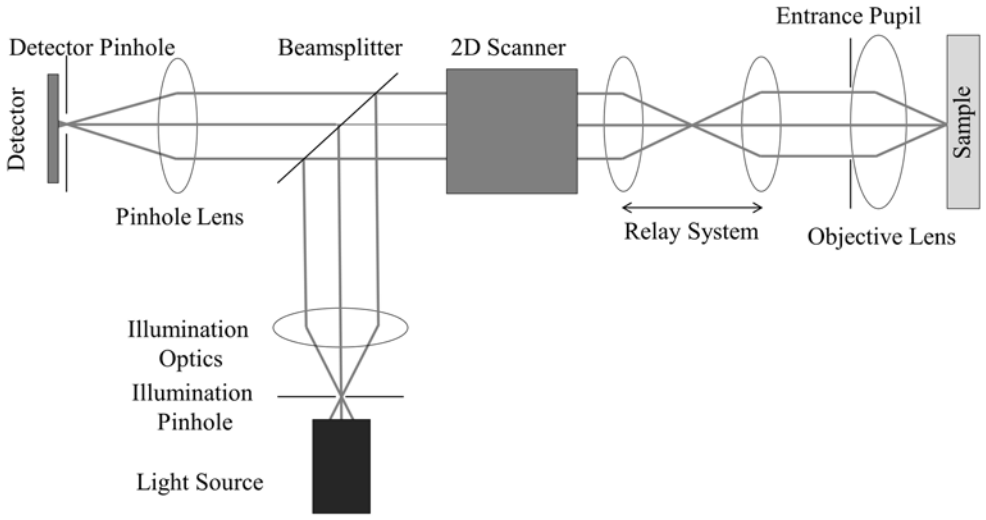

Fig. 1.8 Example of a typical set-up for a confocal microscope

1.3.2.1 Illumination

The light source and related optics in a confocal microscope are responsible for uniformly illuminating the entrance pupil of the objective. The light source must pass through an illumination pinhole. Additional illumination optics are used to ensure the entrance pupil of the objective is uniformly illuminated. One common configuration is to collimate the illumination beam directly after the illumination pinhole as shown in Fig. 1.8.

Lasers are common light sources for confocal microscopes. Lasers are advantageous since they are high power, are stable, and can be focused to a small spot on the sample. If the laser has multiple spatial modes, it has to be spatially filtered. The pinhole used during spatial filtering can simultaneously act as the illumination pinhole. Some incoherent light sources that are used in confocal microscopes are arc lamps and LEDs. The light source should be chosen based on the confocal system design. Some designs are more suited to coherent illumination than incoherent illumination.

1.3.2.2 Beamsplitter

The type of beamsplitter used in a confocal microscope depends on the imaging modality. For reflectance confocal microscopy, a 50/50 beamsplitter may be used. By using this beamsplitter half the illumination light and half the detection light are lost upon reflection. If the reflected signal is low, a beamsplitter with a different transmission/reflection ratio can be used. For instance, a 90/10 beamsplitter, which transmits 90% of the light, is better suited for applications in which the reflected signal is low. Using a high power light source can compensate the loss of 90% of the illumination light upon reflection.

For fluorescence confocal microscopy, a dichroic mirror is typically used as a beamsplitter. The dichroic mirror typically reflects the excitation light and transmits emission light with high efficiency. In addition to a dichroic mirror, an excitation filter and emission filter are needed to further enhance image contrast. The configuration with a dichroic mirror, emission filter and excitation filter is especially desirable for fluorescence microscopy, since the fluorescence signal is often much weaker than the excitation source.

1.3.2.3 Scanner

The easiest way to scan a confocal microscope is to implement stage scanning. This allows the optical system to remain stationary, so there are no complications added to the system. The major disadvantage of stage scanning is that the scanning speed is too slow for real-time imaging.

There are multiple methods that are used to speed up the speed of scanning to video rate. The most common is the use of galvanometer scanners. Galvanometer scanners

consist of the galvanometer, a mirror, and a servo driver that controls the system. The galvanometer mirrors can operate either in closed loop or resonant scanning. The resonant scanning mirrors are advantageous since they yield extremely high scanning speeds. However, at the resonant frequency there is no control over the speed or the angular range of the mirror. The scanning path is often sinusoidal for resonant scanning galvanometers. For this reason closed loop scanning mirrors are used when the scanning speed and scanning range needs to be modified regularly. Closed loop scanning mirrors cannot achieve as high speeds as resonant mirrors, but the scanning path and speed are easily controlled. The scanning mirrors should be optically conjugate to the entrance pupil of the objective.

Polygonal scanners utilize a rotating polygonal optical element with three or more reflective facets. By rotating this polygon at high speeds, the output beam is scanned in a sawtooth pattern. If each facet of the polygon is slightly tilted, the beam position fluctuates from facet to facet.

Acousto-optical scanners utilize a surface acoustic grating to deflect a beam. They are capable of extremely high scan rates, up to 20,000 sweeps per second [19]. However, they provide small angular deflection, limited angle resolution, and low light efficiency. Furthermore, the scan angle is wavelength dependent. Due to the wavelength dependence, this is typically not suitable for fluorescence imaging.

1.3.2.4 Objective Lens

The microscope objective determines the resolution of the confocal microscope. NA of the objective not only determines the resolution of the microscope, but also determines the light collection efficiency and optical sectioning ability. Inherent in microscope

objectives is the tradeoff between NA and field of view (FOV). As the NA increases, FOV and the working distance of the microscope shrink.

For good optical performance, the microscope objective should be well-corrected for aberrations. Since confocal microscopes image details inside tissue, spherical aberration caused by the tissue thickness significantly reduces the resolution. To correct for the spherical aberration induced by the tissue, liquid-immersion objective lenses are often used in confocal microscopy. Objective lenses are typically designed to be telecentric for confocal systems to maintain constant magnification as the microscopes images deep into tissue. The telecentricity also ensures the collection angle for every point on the sample is constant.

1.3.2.5 Pinhole

The pinhole in the confocal microscope determines both the optical sectioning capability and the resolution. The derivation of the PSF for a confocal microscope assumes an infinitely small pinhole. There is an inherent tradeoff in the choice of a pinhole. A smaller pinhole yields better resolution and better optical sectioning; however, the collection efficiency is low. In light starved samples, a small pinhole may not yield an appropriate signal-to-noise ratio (SNR). If the pinhole is increased, the signal is higher, but more light is detected outside the focal point. If the pinhole is increased too much, the system is no longer confocal. A further discussion of the pinhole size on the resolution is found in [20-22].

Typically a pinhole size of 1 Airy unit is chosen, since it has been shown optical sectioning does not considerably improve if the pinhole is smaller [20]. An Airy unit is the radial distance to the first zero of the Airy disk. Since the Airy disk size changes with

wavelength, the optimal pinhole size for different wavelength varies. The performance can be increased for a multi-wavelength confocal microscope if each wavelength has its own pinhole aperture and detector.

1.3.2.6 Relay Systems

The basic requirement of a relay system in confocal imaging is that the aberrations of the relay system should be well controlled and the relay system images the scanning mirrors at the entrance pupil of the objective.

The most straightforward relay system is a 4f system consisting of two identical doublets assembled afocally. The first scan mirror is located at the front focal plane of the first lens, and the second scan mirror is at the rear focal plane of the second lens. The odd aberrations, coma, distortion and lateral chromatic aberration are compensated by the symmetrical configuration. The residual aberrations are spherical aberration, field curvature, astigmatism, and axial chromatic aberration. To address the chromatic aberrations inherent in refractive relay lenses, reflective relay lenses have been developed. Given that the optical properties of a reflective surface only depend on the radius of the mirror, chromatic aberrations disappear.

1.3.2.7 Detector

Common detectors for confocal microscopes are photomultiplier tubes or avalanche photodiodes. Both allow for detection of weak signal with a high SNR. For parallelized confocal systems, the detector may either consist of an array of point detectors (photomultiplier tubes and avalanche photodiodes) or use a high-sensitivity sensor. This high sensitivity sensor be a cooled charge coupled device (CCD), electron multiplying

charge coupled device (EMCCD) or a scientific complementary metal-oxide semiconductor (sCMOS).

1.3.3 Types of Confocal Microscopes

A confocal microscope is inherently a point scanning imaging system. The major issue with a point scanning system is the imaging speed. This section will discuss various scanning methods for improved imaging speed.

1.3.3.1 Point Scanning

The most common confocal system is the point scanning confocal microscope. The illumination pinhole and detection pinhole are confocal on the sample, allowing for improved resolution and optical sectioning capabilities. The excitation light has to be scanned in two dimensions over the entire sample; its major disadvantage is the time it requires to scan over the entire 2D FOV of interest. Figure 1.8 is a typical configuration for a point scanning confocal microscope.

1.3.3.2 Line Scanning

A line scanning confocal microscope scans a focused line across the sample, and the detection pinhole is now a detection slit. Line scanning approach can increase the imaging speed since only one dimension on the sample has to be scanned. In fact, the confocal principle using a slit was described by Goldman in 1940 before Minksy developed the point-scanning confocal microscope [23]. The major limitation is the resolution is only enhanced in the dimension perpendicular to the slit. The resolution parallel to the slit is not enhanced by the slit.

1.3.3.3 Nipkow Disk

Instead of using just a single pinhole for the confocal imaging, a Nipkow disk utilizes a rotating disk with a large number of pinholes. This is essentially a massively parallel confocal microscope. By rotating the disk at high speeds, this allows for extremely fast acquisition of confocal images. Using a Nipkow disk, confocal images can be directly observed with the naked eye.

The Nipkow disk was invented by Nipkow in 1884 as a way to transfer 2D spatial information into a temporal electrical signal by placing a series of holes on a scanning disk. The Nipkow disk was modified by Petran and Hadravsky in 1968 to demonstrate confocal imaging using a tandem-scanning reflected-light microscope (TSRLM) [24]. A major disadvantage of the Nipkow disk confocal microscope is low illumination efficiency, since only a small percentage of the light passes through the small pinholes. If more pinholes are added to the Nipkow disk, there is a greater chance of crosstalk between each pinhole.

While a few methods exist to improve illumination efficiency, the most commonly used is to place a disk with microlenses above the Nipkow disk. The microlens disk is aligned with the Nipkow disk and they rotate together. Each of the microlenses focuses the light onto the corresponding pinhole, as shown in Fig. 1.9. This improves the light efficiency from ~2% to almost 70% [25].

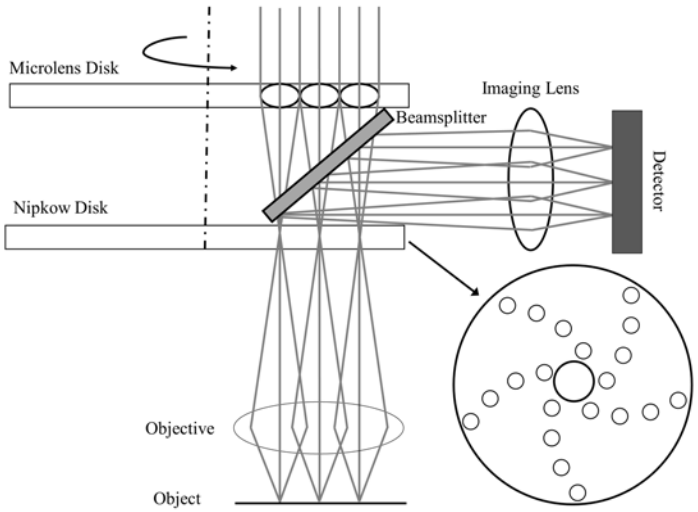

Fig. 1.9 Set-up for a Nipkow disk confocal microscope that utilizes microlenses to increase illumination efficiency.

1.3.3.4 Confocal Microscopes Using Spatial Light Modulators

Another method for parallel confocal scanning is to use a spatial light modulator as the scanning mechanism [26]. A digital micro mirror device (DMD) is one type of spatial light modulator that has been used for confocal imaging. A DMD is a device that consists of hundreds of thousands of micromirrors that are independently switched “on” and “off” at extremely high speeds. The “on” state typically rotates the mirror by 12 degrees, whereas the “off” state rotates the mirror by -12 degrees. Each mirror in the DMD can serve as a pinhole for the illumination and detection path. By temporally switching the micromirrors “on” and “off”, the sample can be scanned at high speed with multiple beams simultaneously. Since multiple beams are typically scanned simultaneously, a CCD or CMOS sensor is typically used to capture the emission signal.

1.3.3.5 Fiber Confocal

Due to the high resolution imaging capability, confocal imaging has many potential clinical applications. However, the size of traditional confocal microscopes makes it problematic to be used for *in vivo* applications. In order to overcome the size constraints,

there are multiple designs that create a small portable confocal microscope that can be used for clinical applications. One such design is to use fibers to create a flexible system with a miniature confocal head. For fiber confocal systems, the fiber acts as both the illumination and detection pinholes. When using fibers for confocal systems, special care should be taken to efficiently couple light into the fibers and to remove Fresnel reflections at the fiber ends, since this acts as background noise for the system. Two designs that utilize fibers for a confocal microscope are single-fiber confocal microscopes and fiber bundle confocal microscopes.

For single-fiber confocal microscopes, light is coupled into a single fiber and the coupled light is output to a miniature confocal imaging probe. The probe typically consists of a method to scan the fiber and a small objective lens [27-34]. The fiber can be mechanically scanned using piezoelectric actuators, electromagnetic actuators, or electromagnetism [35-37]. Another promising method is to uses MEMS mirrors to scan the light output from the fiber. A challenge in using MEMS mirrors is maintaining a small volume for the confocal probe.

Instead of using a single fiber for a small, confocal microscope, a fiber bundle can be used to transfer the light to a small confocal probe [38-39]. Using a fiber bundle eliminates the need for a scanning mechanism in the confocal probe. Since the beam scanning mechanism can be placed at the proximal end of the fiber, a high speed scanning mechanism can be used to scan the beam into individual fibers at the proximal end of the fiber bundle. Each fiber at the distal end acts as the illumination and detection pinhole for the confocal probe. Due to the spacing between the fibers in the fiber bundle, the acquired image is pixelated, which is one limitation of this method. One variant of

this design is to use a DMD to couple light into individual fibers [40]. There are two advantages using a DMD, the scanning pattern does not have to be a raster pattern and multiple fibers can be illuminated simultaneously in parallel, which increases the acquisition speed. Since the entire DMD is illuminated while only a few pixels in the DMD are "on", this design has low light efficiency. Furthermore, if the DMD is not well aligned with the fiber bundle, some pixels on the DMD will image onto the spaces between the fibers in the fiber bundle.

1.3.3.6 Spectral Confocal Imaging

The number of spectral bands measured in a confocal system depends on the number of spectral filters used in the confocal microscope. To distinguish between the signals from multiple fluorescent dyes, multiple optical filters need to be used to differentiate the fluorescent dyes. However, the emission spectra from multiple dyes may begin to overlap. The optical filters may be unable to separate the signal from multiple dyes due to this overlap, which leads to spectral crosstalk in the measurements.

To overcome the limitations of using multiple optical filters in a confocal microscope, a hyperspectral confocal microscope records the emission spectrum for every voxel in the imaged volume by placing a spectrometer after the detection pinhole [41]. Hyperspectral imaging is a natural extension of confocal microscopy, since the emission spectrum is detected after passing through a pinhole, which acts as the entrance pinhole for a spectrometer. The measured spectra are used to identify and distinguish multiple fluorescent dyes in the sample [42]. Even if the emission spectra overlap, the fluorescent dyes can be distinguished through analysis. Hyperspectral imaging has numerous medical applications [43].

1.4 Array Confocal Microscopy

Confocal microscopes are excellent tools for volumetric imaging, since they are capable of noninvasively imaging inside scattering media at high resolution. However, there is a trade-off between the total number of voxels and speed for confocal imaging. Suppose a confocal microscope is used to image a square FOV with dimensions of $L \times L$, where L is in units of length. If the pixel spacing is defined by the Rayleigh resolution Δr of the objective, the resulting image size is $(L/\Delta r) \times (L/\Delta r)$ pixels. If the pixel dwell time is Δt, the total acquisition time is $(L/\Delta r)^2 \Delta t$. Suppose a new microscope objective is designed, such that the resolution is the same, but the FOV increases by a factor of C in each dimension to image a larger area. Therefore, the total number of pixels is $(CL/\Delta r) \times (CL/\Delta r)$ pixels, and the acquisition time is $C^2(L/\Delta r)^2 \Delta t$. Therefore, the acquisition time increases by a factor of C^2. If the dimensions of the FOV are doubled, the total time required to image the area quadruples. Note this is only the time required to image one axial slice of the volume. The time required to image the entire volume is $N_z C^2 (L/\Delta r)^2 \Delta t$, where N_z is the number of axial slices in the volume. This calculation doesn't take into account the time required to move from one axial slice to the next, which is non-zero. If a larger volume needs to be scanned using a confocal microscope at a resolution Δr, then the acquisition time increases significantly and may become prohibitive.

Depending on the size of the volume that needs to be imaged using a confocal microscope, the total acquisition time using a single beam to scan the large volume may be too long. One method to increase the total acquisition time in a confocal microscope is to parallelize the process by designing a system that utilizes N confocal microscopes in

parallel. Instead of scanning the object using a single beam, N beams scan the object in parallel to reduce the total acquisition time by a factor of N. A convenient arrangement for the beams in the object plane is a square array of beams, thus creating an array of confocal microscopes that are each responsible for scanning an area that is $1/N$ of the total FOV. All N images are acquired in parallel and the N images are stitched together to create an image with the desired FOV. The total acquisition time for a large volume is now $N_z C^2 (L/\Delta r)^2 \Delta t/N$. Suppose the acquisition time for imaging a volume of interest takes 9 hours using a single beam. If 9 beams are used to scan every slice in the volume, then the total acquisition time to image the same volume reduces to 1 hour, in theory. This significant decrease in acquisition time has the potential to lead to the study of larger volumes of interest, since acquisition time is not as big of a factor.

The above example described the increase in acquisition time resulting from increasing the FOV of a microscope objective by a factor of C. For objectives there is a tradeoff between NA and FOV. Designing an objective that increases the FOV of an objective while maintaining the resolution may be an extremely difficult task depending on the size of the resulting FOV. To significantly increase the FOV of an objective without changing the resolution, the objective typically has to grow in size, complexity, and cost. While there is increased difficulty and cost in designing this objective, there is a significant decrease in the overall acquisition time for large volumes. This significant speed up in acquisition time may yield benefits that outweigh the increased cost of the system.

For an infinity conjugate objective, an incident oblique collimated beam focuses in the sample plane a distance from the optical axis defined by the angle of incidence of the

beam. Therefore, to create an array of N beams in the sample plane, N collimated beams are incident on the objective at N distinct angles. These N angles map to N different positions in the sample plane. When using a single objective for an array confocal microscope, a system needs to be designed such that N collimated beams enter a microscope objective at some designed angles to focus the N beams onto the sample plane in some designed pattern.

When using a single objective for an array confocal microscope, the tradeoff between NA and FOV needs to be considered. This tradeoff limits the overall resolution and FOV in an optical system, since there are always manufacturing considerations and a budget for any optical system. Instead of designing an array confocal microscope using a single objective, an array of objectives can be used to create an array confocal microscope. Each confocal microscope in this configuration has its own objective and is responsible for scanning a sub-FOV. One design for such a system is to design an array of miniature objectives that are designed for on-axis performance. The array of miniature objectives is spatially scanned by moving the entire array assembly. The images from all objectives are then combined to create one larger FOV image. Since the array of objective is spatially scanned, the FOV isn't limited by the FOV of a single objective; instead the FOV is limited by the number of objectives in the array. By using this configuration, the tradeoff between FOV and NA is overcome.

For an array confocal microscope there is the tradeoff between NA, FOV, speed and cost. Using multiple confocal microscopes in parallel decreases the total acquisition time required for scanning. However, the optical system still has to be designed with the four tradeoffs in mind. Chapter 2 explores an array of miniature objectives to create an array

confocal microscope that overcomes the tradeoff between NA and FOV; however, the speed is limited in this design. Chapter 4 demonstrates an array confocal microscope in which a microscope objective is designed to have a high NA and a large FOV simultaneously, however, the cost of this system increases. A major challenge in any system is to balance these 4 competing tradeoffs.

For an array confocal microscope, *N* beams are used. Instead of using *N* lasers to create an array of *N* beams, it is convenient to utilize a method to split one laser beam into an array of *N* beams. While there are several ways to split one laser beam into several outgoing beams, diffraction gratings are advantageous since they have relatively high efficiencies and diffract beams at well-defined angles. Therefore, beam-splitting diffraction gratings are ideal components for array confocal microscopes.

1.5 Outline for Dissertation

This dissertation explores two designs for an array confocal microscope. The first design utilizes an array of miniature objectives with a high NA to create a confocal microscope with both a large FOV and high NA. This design overcomes the trade-off between NA and FOV in a microscope. If the FOV of the objective needs to be increased while keeping the NA constant, generally the size of the objective also has to increase. This is what is done in the second design. The second design utilizes a large microscope objective designed to have a high NA, large FOV, and large working distance. In order to speed up acquisition of the large FOV, an array of *N* beams is scanned across the samples to reduce the acquisition speed by at least a factor of *N*. Both of these designs utilize beam-splitting phase gratings to create an array of beams at the sample plane. The remainder of this dissertation is devoted to exploring aspects of beam-splitting phase

gratings. Continuous phase gratings yield the highest efficiency, so a method to optimize continuous phase beam-splitting gratings that are less sensitive to fabrication errors is developed. Typically, phase gratings only perform as designed for a single wavelength; so achromatic phase gratings are explored as a way to design a grating to perform as designed for two or more wavelengths. Through the analysis of the achromatic phase gratings in the presence of fabrication errors, it is proved that phase profiles with constant first spatial derivatives are the optimal designs. Even though achromatic phase gratings perform as designed at more than one wavelength, the diffraction angles of those two wavelengths are different. In the absence of other optics, the beams at the two wavelengths will not overlap. An achromatic scale-invariant phase grating is designed, which has the designed grating performance at two wavelengths, while also diffracting at the same angle at two wavelengths.

Chapter 1 introduces confocal microscopes, as well as introducing the concept of an array confocal microscope. Chapter 2 presents a design for an array confocal microscope consisting of an array of miniature objectives to expand the FOV without limiting the NA of the objective. Chapter 3 proposes a new method to optimize beam-splitting phase gratings with reduced sensitivity to fabrication errors. Chapter 4 demonstrates a design with a large objective and working distance that scans an array of beams to significantly speed up the acquisition speed of a confocal microscope. Chapter 5 analyzes achromatic beam-splitting phase gratings. Chapter 6 designs a scale-invariant achromatic phase grating, which are phase gratings that not only have the designed performance at exactly two wavelengths, but also equal diffraction angles at those two wavelengths. Chapter 7

describes possible designs for an achromatic confocal microscope. Chapter 8 concludes the dissertation with a summary of the current research and future directions.

2. ARRAY CONFOCAL MICROSCOPE USING ARRAY OF MINIATURE OBJECTIVES

2.1 Introduction

Due to its sub-cellular resolution, confocal microscopy is an excellent tool for the detection and diagnoses of many types of cancers, however, there are several limitations that prevent this technology from being used clinically. The three main limitations of confocal microscopy that prevents clinical use are the field of view (FOV), size, and speed.

In any microscope, there is an inherent trade-off between the numerical aperture of the objective (NA) and the field of view (FOV). As the numerical aperture of the objective increases, the FOV correspondingly decreases. For example, in PLN objectives a numerical aperture of 0.25 corresponds to a FOV of 2.2 mm. If the numerical aperture increases to 0.65, the FOV decreases to 0.55 mm [44]. In order to increase the FOV of an objective at a given NA, a larger objective is needed. A larger objective is both more difficult to fabricate and more expensive.

Speed is also a major limitation of confocal microscopy. There are several fast confocal systems, such as line scanning or Nipkow disk. Line scanning sacrifices resolution along the scanning line, and the Nipkow disk is too large to be useful in clinical applications. Confocal microscopes are usually rather large and bulky. Their size prevents their use in clinical settings, where a small, portable probe is usually required. Currently, hand-held confocal microscopes have been demonstrated, but their FOV is generally very small, usually much less than 0.5 mm [27-40].

An array of microlenses has been used for point scanning fluorescence imaging to create gigapixel images [45]. The sample is scanned laterally in two directions to cover the sub-FOV of each microlens. All images are stitched together to create one gigapixel image. The concept of using a microlens array for confocal microscopy has also been demonstrated [46-48]. A FOV of 5.5 mm × 5.5 mm with an array of 0.41 NA microlenses was demonstrated [48]. Spherical aberration prevented the diffraction-limited performance since only a single lens was used for the objective array. Instead of utilizing an array of microlenses, a chromatic confocal microscope was demonstrated that used a pinhole array in conjunction with a hyperchromatic objective for confocal microscopy without axial scanning [49]. The hyperchromatic objective focuses wavelengths at different axial positions, so the axial position of the sample is determined without any axial scanning. However, the efficiency through the pinhole array is very low. Due to the size and configurations of these array confocal microscopes, they are only suited for stage scanning.

This chapter explores a design that uses an array of miniature objectives to extend the FOV of a confocal microscope without reducing the NA. Instead of utilizing stage scanning to scan the sample over a sub-FOV for each objective, the system is designed for potential use with hand-held applications. The objective array is scanned laterally, while the sample remains stationary. Furthermore, an objective that utilizes more than a single element is designed for diffraction-limited performance. A method to fabricate the microscope is proposed, however, due to sensitivity to fabrication tolerances, the system is not built.

2.2 Concept and Design of Array Confocal Fluorescence Microscope

The proposed array confocal fluorescence microscope (ACFM) is a small, portable hand-held device that simultaneously has a high NA and a large FOV, which is currently contradictory in microscopy. The design allows it to image large 3D volumes faster than conventional confocal microscopes over a large FOV.

Instead of using a single high NA objective to scan a large FOV, a miniature objective array is used in order to significantly increase the FOV, while maintaining a high NA. Each miniature objective simultaneously scans a sub-FOV, represented as the small squares in Fig. 2.1(a). The overall FOV of the system is the sub-FOV of each miniature objectives times the number of objectives used. Therefore, the FOV is easily scalable by increasing the number of objectives in the array. Note that the array is not limited to a square array, it can be any size or geometry required by the user.

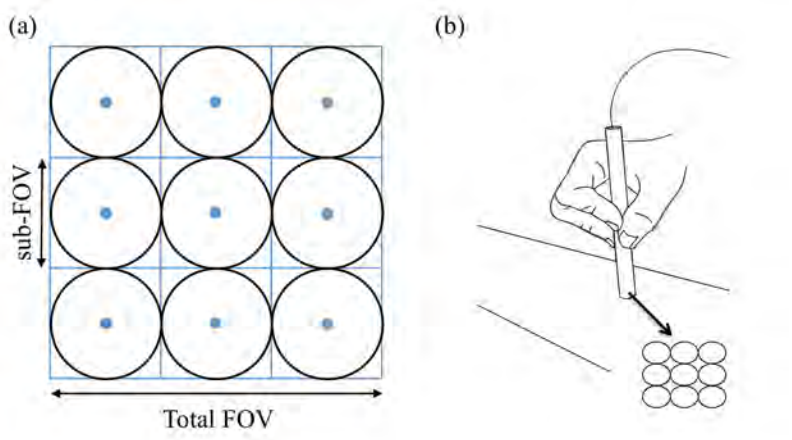

Fig. 2.1 (a) Concept for the miniature objective array. Every circle represents a high NA objective. Each objective scans a small area represented by the small squares. The total FOV of the objective array is the sub-FOV of each objective times the total number of objectives. (b) The concept for the miniature objective array houses it in a small, portable device that will easily fit in a hand.

The miniature objective array is housed inside a small, portable confocal head. This portable confocal head should contain both the miniature objective array and a scanning mechanism. An ideal representation of the confocal head is shown in Fig. 2.1(b). Due to

its small size and portable nature, the ACFM is an ideal candidate for clinical applications. It has potential applications in the diagnoses and detection of skin cancer, oral cancer, and many more diseases [50,51].

2.2.1 Illumination and Detection Optics

The system is designed such that the excitation and emission light travels to and from the confocal head through a fiber array. Therefore, the illumination and detection optics are housed separately from the small, portable head. Since the illumination and detection optics are easily modified, this system is easily adapted to other advanced imaging modalities such as multiphoton microscopy or hyperspectral confocal microscopy.

The design for the illumination and detection optics is shown in Fig. 2.2(a). Collimated light from a 488 nm argon-ion laser is expanded using a beam expander and is incident on a beam-splitting diffraction grating. This diffraction grating splits the incident beam into N equal energy outgoing beams, where N is the desired number of outgoing beams. Each of the N beams is reflected off a dichroic mirror, where the light from each beam is coupled into a $1\times N$ fiber array. Each of the N beams travels through the fiber array, exiting through the opposite end of the fiber array in a $N^{1/2}\times N^{1/2}$ pattern. The miniature objective array focuses the N beams onto a sample, where each of the N spots on the sample are simultaneously excited, as shown in Fig. 2.2(b). The sample fluoresces at each of these N spots, travels back through the miniature objective array, and then travels through the fiber array. Note that the diameters of the fiber act as the illumination and detection pinholes for the confocal system. The emitted light exits the fiber array, travels through the dichroic mirror and is detected on a detector. This detector

can be a high-speed sCMOS, a CCD detector, or it can be a multi-channel photo-multiplier tube to take advantage of its increased speed.

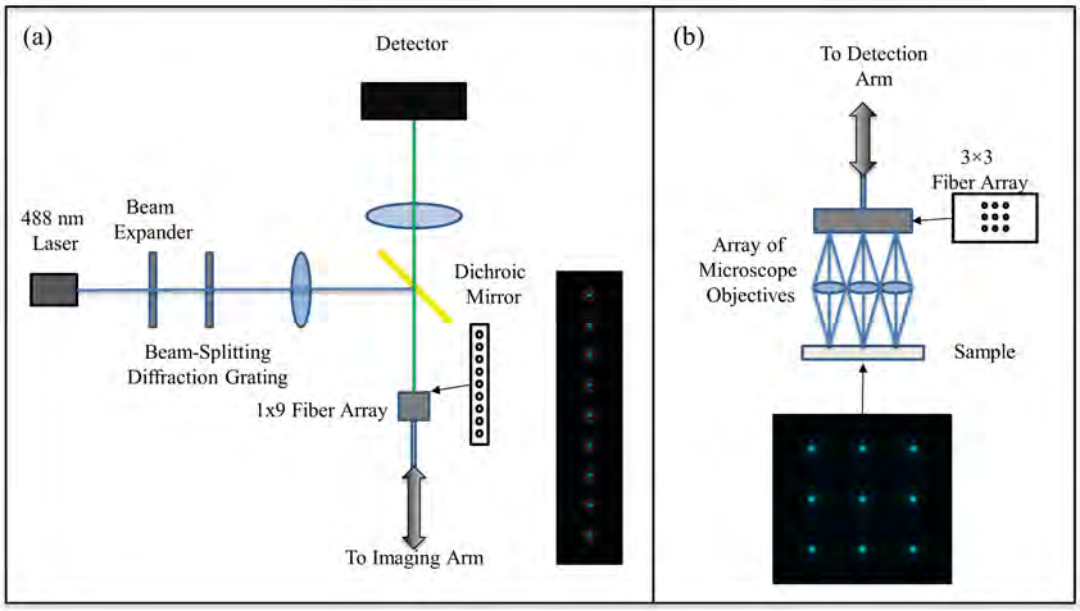

Fig. 2.2 The experimental set-up used to demonstrate the array confocal fluorescence microscope.

2.2.2 Experimental Validation

The design of the illumination and detection system is built and experimentally tested. Instead of using an array of miniature objectives, a single objective is used to simulate the array of miniature objectives. This section explores and tests the optics used in the illumination and detection system.

A 488 nm argon-ion laser expands its size by 5× using a beam expander; this expanded beam is then incident on a continuous phase beam-splitting diffraction grating. This grating has a theoretical efficiency of 94.5%. The design, fabrication and testing of this beam-splitting grating is explained in detail in Chapter 3.

2.2.3 Fiber Array

The 9 outgoing beams from the beam-splitting diffraction grating are coupled to a fiber array with 1×9 fibers equally spaced on one end and 3×3 fibers equally spaced on the

other end as shown in Fig. 2.3. The fiber array that was used contains photonics crystal fibers with a core diameter of 20 μm, a NA of 0.08, and a fiber spacing of 500 μm. Each outgoing beam from the diffraction grating is approximately 13 μm when focused on the end of the fiber array. The coupling efficiency through the fiber array was measured to be 65.5 ± 2.9%, with a maximum energy difference between beams of 15 ± 2%. Therefore, the combination of the beam-splitting diffraction grating and the fiber array is an efficient method of splitting *N* beams into a geometry that is useful for the ACFM.

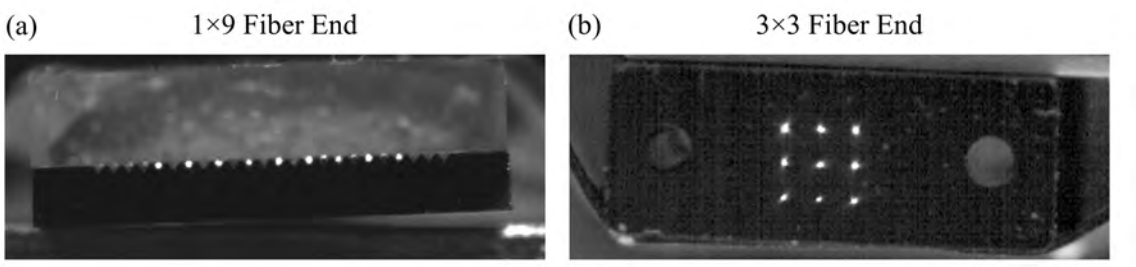

Fig. 2.3 Photo of the (a) 1×9 end of the fiber array and (b) photo of the 3×3 end of the fiber array.

2.2.4 Results with a Single Objective

To test the array concept, a single 0.5 NA microscope objective is used instead of the array of miniature objectives, as shown in Fig. 2.4(a). A mouse kidney section stained with Alexa Fluor 488 WGA is translated 200 μm in each direction using a translation stage. The fluorescence is detected using an sCMOS sensor after passing through the detection optics. Note the fiber array acts as the illumination and detection pinhole for this confocal system. In the same amount of time it took to scan a 200 μm × 200 μm area, an approximately 600 μm × 600 μm area is imaged, as shown in Fig. 2.4(b). Note that each sub-image in Fig. 2.4(b) corresponds to the signal obtained from each of the 9 fibers and each image was obtained at the same time. Therefore, a 9 times increase in the FOV is obtained in the same amount of time it would take to scan a single image. Since each of

the images contains a small overlapping region, these 9 images can be stitched together in order to output a single image large FOV image.

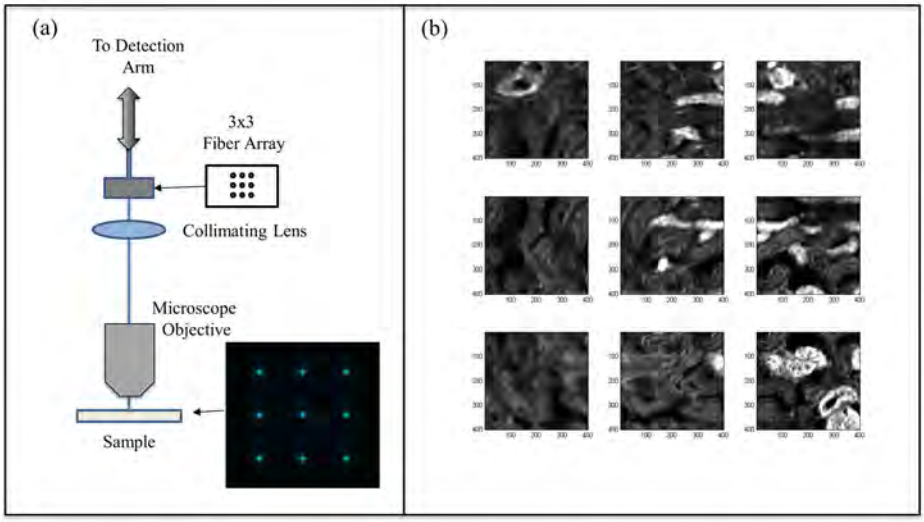

Fig. 2.4 (a) A single objective is used to focused the array of 9 beams onto the sample. (b) This image of a mouse kidney section stained with Alexa Fluor 488 WGA was obtained by using a single NA 0.5 objective and a 3×3 fiber array. The sample was moved 200 µm in each direction. By using the array concept an approximately 600 µm × 600 µm area was obtained in the same amount of time it would take to scan a 200 µm × 200 µm area.

Since this system is multi-channel, the crosstalk between channels is measured. To perform this measurement, the excitation light is sent through a single fiber and the emitted light is recorded from all 9 fibers. The results are shown in Fig. 2.5(a) and 2.5(b). As shown in Fig. 2.5(a), a signal is only obtained from the fiber in which the excited light is sent through. Fig. 2.5(b) shows the values of the pixels of the image sorted by the intensity of the pixels in the middle image. Note that a rise in the intensity of the middle image does not produce any increase in the intensity of any other image. Therefore, there is no detectable cross talk in this system. Note that the separation between excitation spots on the sample is approximately 200 µm with this set-up. When the miniature objective array is used, the spacing between excitation spots is going to be 500 µm. Due to this increased distance, there will be even less chance of crosstalk between channels using the miniature objective array.

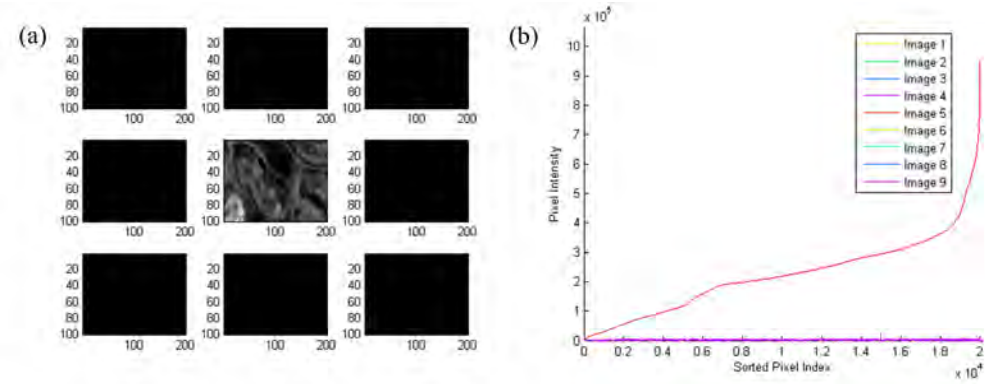

Fig. 2.5 (a) The image obtained when light is only sent through a single fiber. (b) The pixels are sorted based on the intensity of the central image. Note that an increase in the intensity of the middle image does not yield an increase in the intensity in the other channels. Thus, there is no detectable crosstalk in this system.

From these experiments, the combination of a beam-splitting diffraction grating and a fiber array is demonstrated to be an efficient beam-splitting method for the ACFM. Furthermore, it is demonstrated that the fiber array efficiently transfers an array of 1×9 beams into an array of 3×3 beams. The 3×3 array of beams simultaneously excites the sample, and the fluorescence from each of these 9 beams is simultaneously measured on a detector with negligible crosstalk.

2.3 Design of the Miniature Objective Array

The excitation system shown in Fig. 2.2(a) is going to be used to transfer the 9 equal energy beams to the array of miniature objectives. The concept for the array of miniature objectives is shown in Fig. 2.6. Figure 2.6(a) shows the concept for a hand-held array confocal microscope. The array of miniature objectives are housed in a barrel, and this barrel is going to be translated in 3-dimensions to act as a scanning mechanism. The array of miniature objectives is enclosed in the barrel; a glass plate is placed on the end of the barrel to protect the array of miniature objectives. A conceptual diagram inside the barrel is shown in Fig. 2.6(b). The light from the fiber array exits in a 3×3 array, where the centers of each fiber are aligned with the optical axes of an array of miniature objectives.

Each of these miniature objectives focuses the 9 beams into a 3×3 array. Note the diameter of the fibers act as the illumination and detection pinhole for the confocal system.

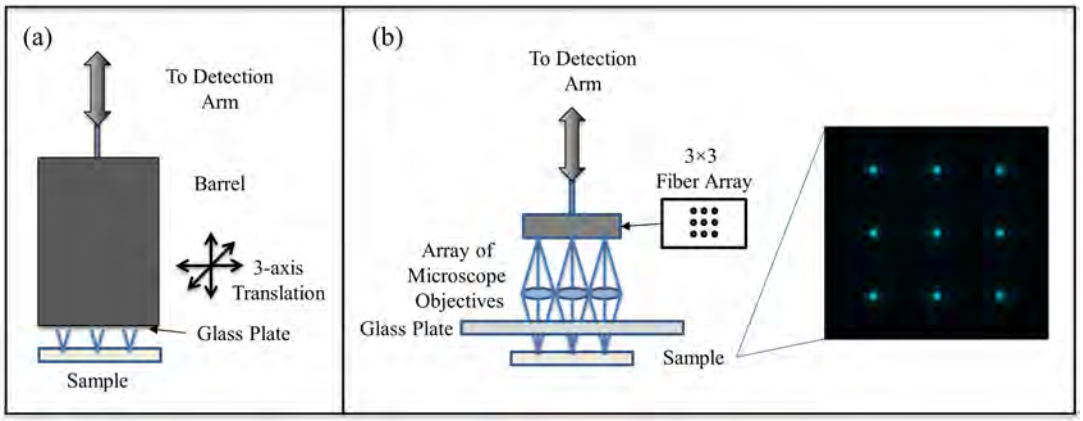

Fig. 2.6 (a) Barrel that houses the array of miniature objective. (b) Concept for array of miniature objectives.

The miniature objective array is designed to have the following specifications. Each miniature objective has a NA of 0.5 with a spacing of 0.5 mm between the centers of each objective. The objectives are optimized for an excitation wavelength of 488 nm and an emission wavelength of 519 nm. Since the objective array is going to be scanned along with the fibers, this system works entirely on-axis. Therefore, there is no loss in image quality at the edges of the field of view. From the manufacturer of the fiber array, the tolerance in the positioning of the 3×3 array of fibers is ±20 μm in the horizontal and vertical directions of the fiber. Therefore, the design for the miniature objective is designed with a field of ±20 μm to optimize the performance of the objective over this error range.

The fabrication method for the miniature objective array is single point diamond turning. By using single point diamond turning, aspheric surfaces can be used, which allows for increased imaging performance. Additionally, another advantage of using

single point diamond turning is the alignment fixtures can be fabricated directly on the lens array itself allowing higher precision during the alignment and assembly process.

2.4 Mechanical Design

This section describes a design for the mechanical design and scanning method for the miniature objective array confocal microscope. Figure 2.7(a) shows the mechanical design for the barrel that positions the fiber array with respect to the miniature objective array. The barrel consists of two parts, the fiber array holder and the miniature objective array holder. There are four sources of error in the alignment of the fiber array with respect to each miniature objective. First, the centers of the fibers in the fiber array must be rotated to align with the miniature objective array. Then there is the potential for translational misalignment in all three spatial dimensions. Figure 2.7(b) shows a method to lock the rotational axis of the miniature objective array. A rectangular notch is cut into the barrel and a matching rectangular extrusion is fabricated on the plastic substrate the miniature objective array is fabricated on. The fiber array holder now has the freedom to rotate about the central axis of the barrel. Three holes are drilled 120 degrees about into the barrel and tapped. Three holes are also drilled into the fiber array holder in locations with a larger diameter than the screw diameter, as shown in Fig. 2.7(c). The larger diameter allows a slight adjustment to the rotation angle of the fiber array holder, which acts as a compensation method to better align the fiber array holder. Three holes are also drilled into the sides of the barrel and tapped, such that set screws can be screwed into the side. These set screws are used to adjust the lateral position of the fiber array holder. Finally, there is no compensation method for the axial position of the fiber array to the miniature objective array. This distance must be machined to some pre-defined tolerance.

This tolerance is taken into account during the lens design section of this chapter. A prototype of the barrel is 3D printed using a Makerbot 3D printer, and a picture of this prototype is shown in Fig. 2.7(d). To demonstrate the size of this barrel, a penny is placed next to the mount. A view from the bottom of the barrel in Fig. 2.7(e) shows light exiting the 3×3 fiber array in the barrel. Note there is no lens array in this prototype, it is just demonstrating the design of the barrel and the placement of the fiber array in the barrel.

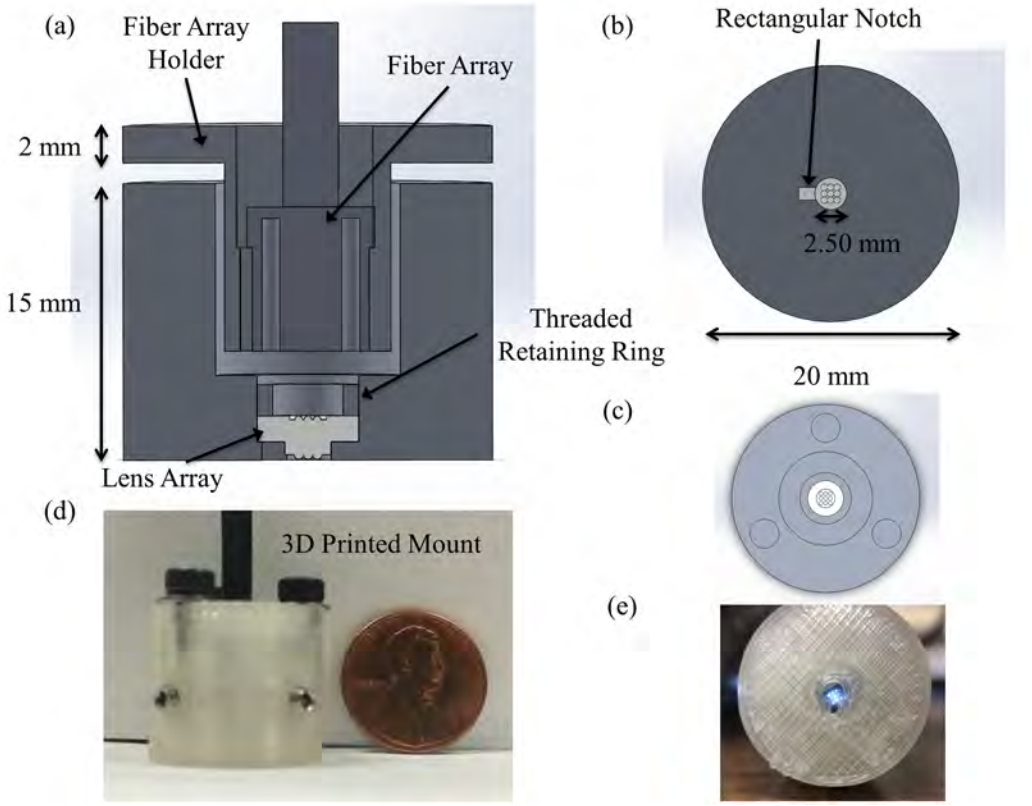

Fig. 2.7 (a) Cross section of barrel that holds the fiber array and minaiture objective array. (b) Bottom view and (c) top view of barrel. Picture of (d) side-view and (e) bottom view of 3D printed prototype when light is sent through fiber array.

To scan the miniature objective array, a 2-axis scanner is placed on the barrel to move the barrel as an assembly. To protect the lens array from damage, a glass plate is placed in between the barrel and the sample. A diagram of the scanning in relation to this glass plate is shown in Fig. 2.8(a). Figure 2.8(b) shows the bottom view of the barrel through the glass plate. The glass plate has a diameter of 25 mm and a thickness of 0.2 mm.

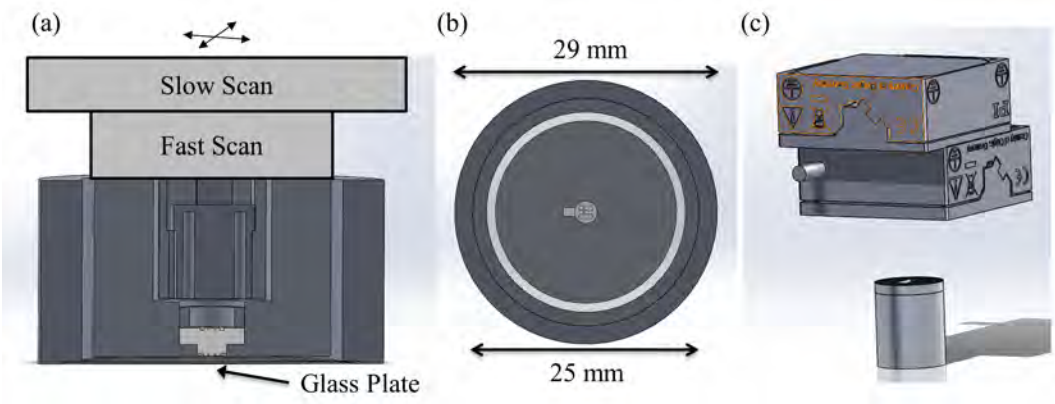

Fig. 2.8 (a) Side view of concept for mechanical mount, glss plate, and scanning directions. (b) Bottom view of barrel and glass plate. (c) Potential translation stage and barrel.

Piezotubes are often used for fiber scanning confocal microscopes, however, this system needs to scan 0.5 mm in each direction, which is much larger than the maximum *xy* deflection for piezotubes. For instance, PI's PT230.94 piezo tube has a maximum *xy* deflection of ±35 μm. A translation stage with a range larger than 0.5 μm and resolution < 0.5 μm is required. Additionally, this translation stage should be both fast and compact. As the resolution of the translation stage increases, the travel range and the speed typically decrease and/or the size of the translation stage grows considerably. The M663 precision positioning stage is determined to be the best trade-off between all these design requirements and the cost [52]. This translation stage has 18 mm travel range with a ±0.5 μm bidirectional repeatability. The maximum velocity of the stage is 250 mm/s. The size of the M663 positioning stage in relation to the barrel is shown in Fig. 2.8(c). A mount would be designed to attach the barrel to the 2-axis stage, in order to translate the array of miniature confocal microscopes laterally to scan the sample.

To determine the scanning speed of this system, assume the stage has an infinite acceleration, so it immediately reaches its maximum velocity to calculate the minimum time to scan over an 0.5 mm × 0.5 mm area. The maximum driving frequency of the stage is 500 Hz in the fast axis for a 0.5 mm travel range. Assume the pixel spacing in the

image is 0.5 μm, therefore the image is 1000×1000 pixels. Therefore, the absolute minimum time to scan a 0.5 mm × 0.5 mm using this scanner is 2 s. In actuality, the stage has a finite acceleration, which determines the time needed for the stage to reach a 250 mm/s velocity. Therefore, in practice the scanning speed is going to be longer than 2 s. This scanning speed may be problematic for clinical applications, especially if it is going to be used as a hand-held scanner. The object may move significantly during this 2 s window, which is going to add motion blur to the acquired images and may make them unusable in practice.

2.5 Single Lens Objective Design

Ideally, the array of miniature objectives would be a single element, since a single element is much easier to align than a several element system. To test whether a single element design is feasible, a 0.5 NA objective with a diameter less than 0.5 mm is designed using PMMA, since PMMA is easily diamond-turned. A 0.2 mm thick glass plate is placed between the objective and the sample. The sample is assumed to have the index of water, since tissue has a similar index to water. A single element design for this objective is shown in Fig. 2.9(a). The objective is not positioned directly on the glass plate, since the weight from the array of miniature objective has the potential to warp the shape of the glass plate due to the weight of the array of objectives. The lens looks like a cylinder, since the thickness is 2 mm and the diameter is 0.5 mm. Significant warping in the shape of the objectives may occur during the diamond turning process if the thickness is less than 2 mm. Each surface for the single element lens has an aspheric coefficient. The distance between the glass plate and the image plane is 3 μm. The performance for this lens is poor, as shown in the MTF in Fig. 2.9(b) and the spot diagram in Fig. 2.9(c).

There is a significant drop in the MTF of this system, and the spot size is greater than the diffraction limited spot-size for all 5 fields. Additionally, this poor performance doesn't even take into account fabrication errors during the fabrication process. Therefore, a single element design is not possible due to two major reasons: spherical aberration and axial chromatic aberration. Additional lens elements are needed to reduce these aberrations.

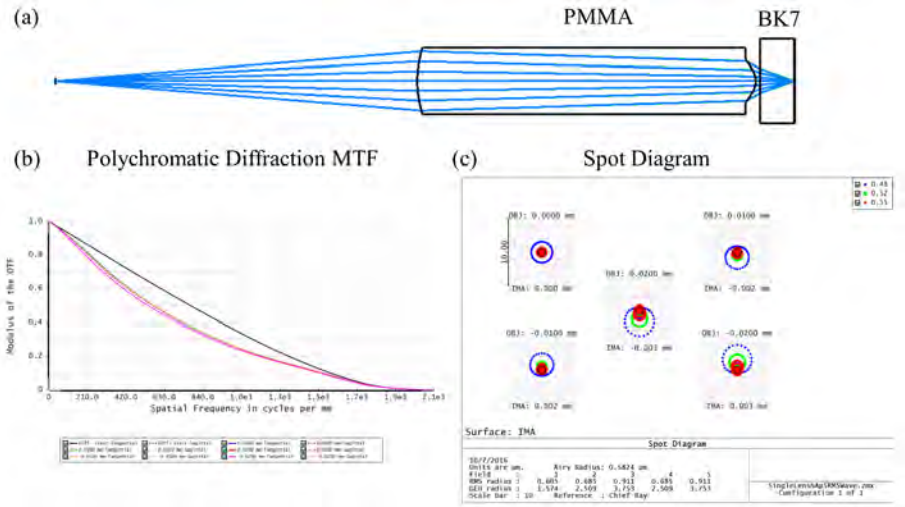

Fig. 2.9 (a) Optical design for miniature objective using a single lens. (b) MTF for objective. (c) Spot diagram for objective.

2.6 Three Lens Design

To effectively remove the spherical and axial chromatic aberration, the miniature objective array is designed using three elements. Figure 2.10(a) shows the optical layout for this three lens design. Two optical plastics are used, PMMA and polystyrene. PMMA has n_d = 1.49 and v_d = 57.4, and polystyrene has n_d = 1.59 and v_d = 30.9. The combination of these two plastics allows for the reduction of chromatic aberration. Due to the small diameters of the lenses and the requirement that the thicknesses of the lenses are greater than or equal to 2 mm, the lenses appear to be rod-like. Due to this rod-like behavior, the design is fairly limited. The design is as follows: the first element collimates the light

from the fiber. The second element is a negative lens that helps to correct the chromatic aberration, and the third element focuses the light at the designed 0.5 NA. The index in image space is water. The majority of the optical power in the objective is in the last surface of the third element. Note all lens surfaces are parameterized using even aspheres. A magnified view of the last lens surface is shown in Fig. 2.10(b).

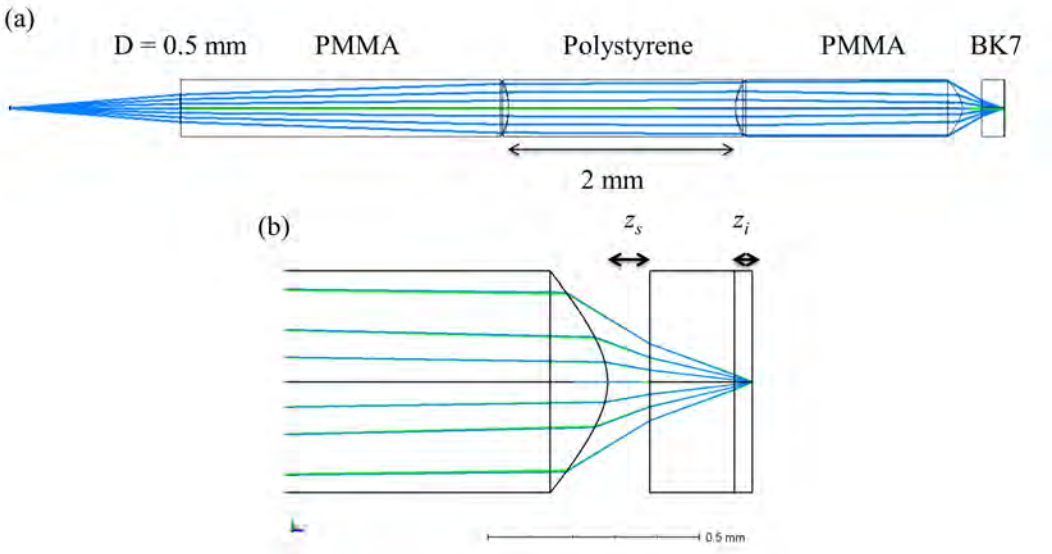

Fig. 2.10 (a) Optical design of three element miniature objective. (b) Magnified view of the last lens surface and glass plate of the objective.

There are two major sources of error that are considered when optimizing these lenses: decentering of surfaces within a single lens and decentering of lenses with respect to each other. The first error will arise during the diamond turning process since only one side of the lens is fabricated at a time. One side of the lens has to be diamond turned, then after that side is finished; it is flipped upside down to be diamond turned on the other side. When the lens is flipped, the position of the tool with respect to the optical axis of the first diamond turned surface cannot be perfectly aligned. Another major source of error is the decentering of elements with respect to each other. Even though alignment features are added to the lens itself, there is still some error in the alignment. This objective is designed with these tolerances in mind; 21 configurations are used to simulate decenters

between elements up to 10 µm and errors in the separation of elements up to 100 µm. The fields in Zemax are defined as object space telecentric with an object height of ±20 µm to simulate the tolerance in the 2D positioning of fibers in the fiber array. Since this objective needs to be scanned in 3D, multiple configurations are used to scan the focused beam in the image plane. The objective lens is translated axially behind the glass window to scan the beam in the image plane. In Zemax, the objective is modeled using the multi-configuration editor to simulate decentering between elements and surfaces. The system is optimized to minimize the merit function in all configurations. After optimization, an axial scan of 75 µm of the lens behind the glass results in an image axial scan of 112 µm in image space. Figure 2.11 shows the MTF at the four specified axial positions. Table 2.1 shows the RMS spot size for the specified axial positions.

Table 2.1 RMS Spot Size for Specified Axial Positions

z_s	z_i	RMS Spot
50	112	0.550
75	77	0.424
100	42	0.396
125	0.01	0.482

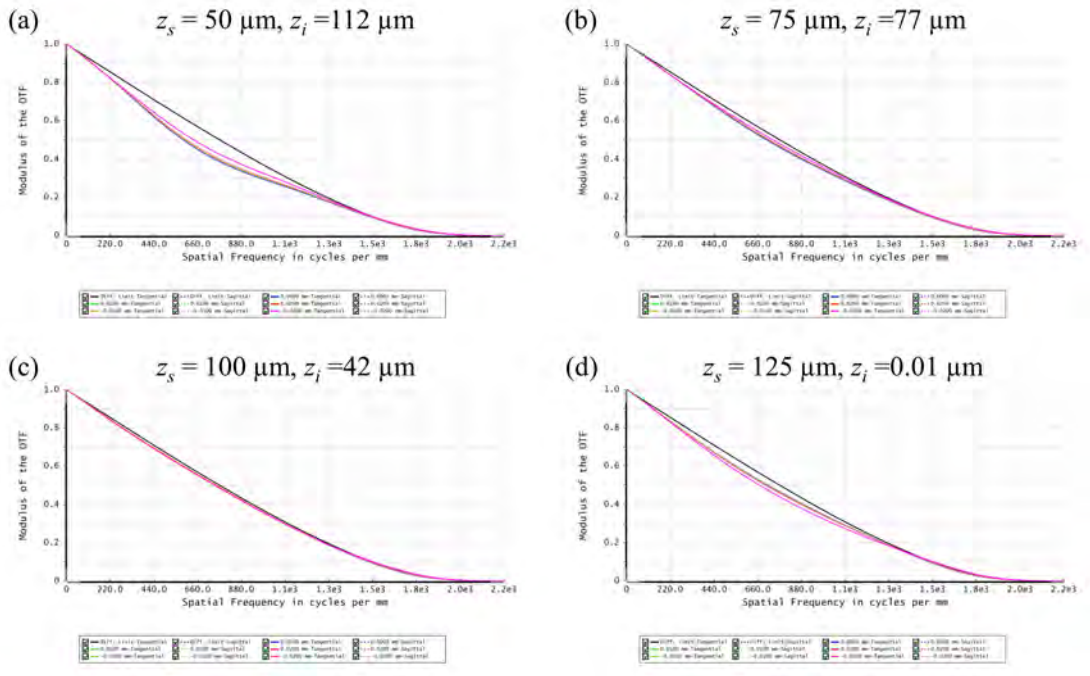

Fig. 2.11 (a) – (d) MTF of three lens objective at specified axial positions.

2.7 Tolerancing the Three Lens Objective

This section explores the tolerancing of the decentering of elements and surfaces of the lenses in the objective. To study the effect of just the decentering between elements and surfaces, the error in the positioning of the fiber array is ignored in this analysis. For the tolerance analysis, there is a single field specified with an object height of 0 μm. The objective is only analyzed for one axial position: z_s = 100 μm, z_i = 42 μm. The last surface is used as a focus compensator during the tolerancing. The test wavelength is the wavelength of the excitation light 488 nm. All elements and surfaces have a decenter range of ±δ, the thickness of all elements has a range of ±2δ, except for the thickness between elements 1 and 2, which has a range of 0 - 2δ to prevent the lenses from passing through each other in the design. A Monte Carlo analysis is used to simulate perturbations in all lens elements simultaneously. At every Monte Carlo cycle, each surface has a perturbation in the thickness and decenter based on the value of defined perturbation range. The criterion for the Monte Carlo tolerance analysis is RMS wavefront. A summary of the Monte Carlo analysis results for different values of δ is shown in the next few sub-sections.

2.7.1 δ = 10 μm

Since the objective is optimized with configurations of decenters up to 10 μm, a Monte Carlo analysis with δ = 10 μm is performed. The two surfaces that are the worst offenders are decenters of the two surfaces in the last objective element, which makes intuitive sense. The last lens surface in the objective has the majority of optical power, so a decenter in that surface should have a significant degradation in the performance. The Monte Carlo analysis performs perturbations in 1,000 samples. From these 1,000 samples

the best RMS wavefront is 0.027 and the mean RMS wavefront is 0.174. Based off the Monte Carlo analysis 50% of the samples have RMS wavefronts less than 0.145, and 90% of the samples have RMS wavefronts less than 0.286. Figure 2.12 shows the spot diagram and the MTF for a sample close to the 90% and 50% performance. Note the performance for both instances is very poor. Based off the spot diagram, the decentering of the lens surfaces and elements introduces coma into the system. This results in performance that is very far from the diffraction limit. Therefore, a tolerance of 10 μm is too large for the decenter of lens elements and surfaces.

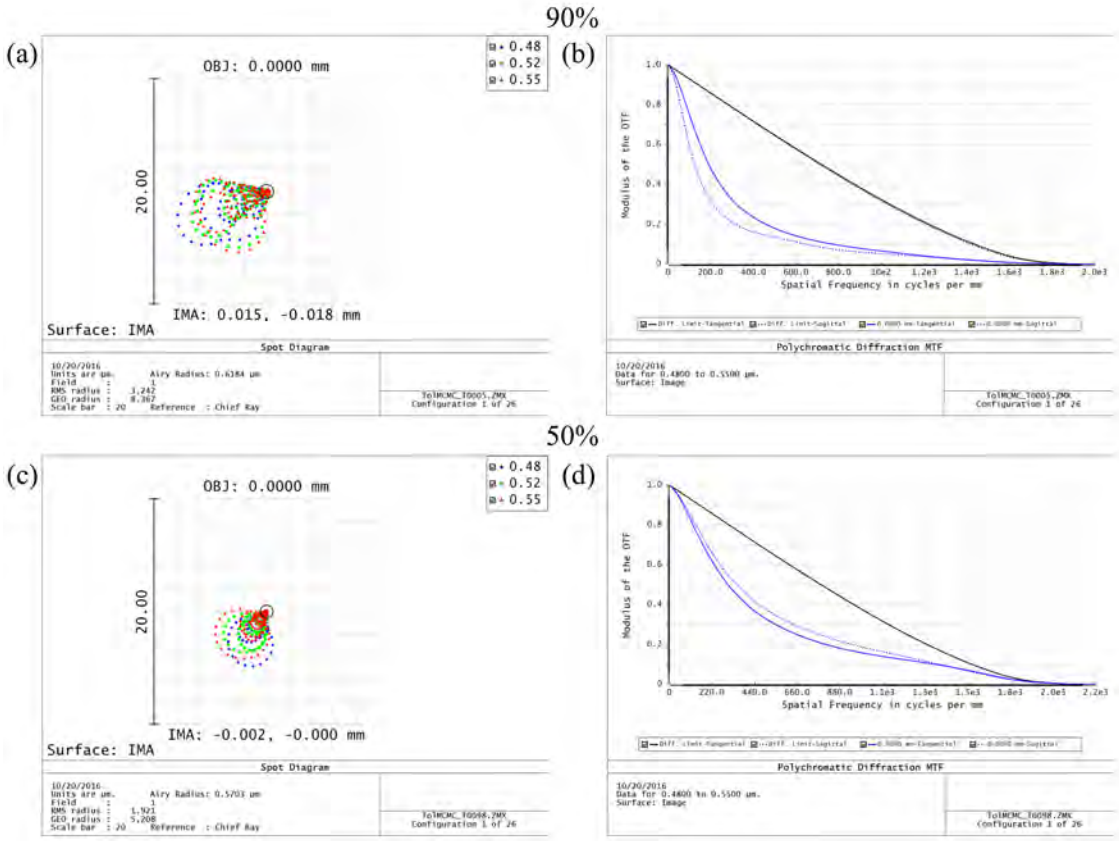

Fig. 2.12 (a) Spot diagram and (b) MTF for sample close to the 90% performance when δ = 10 μm. (c) Spot diagram and (d) MTF for sample close to the 50% performance.

2.7.2 δ = 5 μm

Since δ = 10 μm is too large of a tolerance for the decentering of elements, a Monte Carlo analysis with δ = 5 μm is performed. Once again, 1000 samples are analyzed in the

Monte Carlo simulation. For the 1,000 samples the best RMS wavefront is 0.022 and the mean RMS wavefront is 0.069. Based off the Monte Carlo analysis there's a 50% probability that the RMS wavefront is less than 0.065, and a 90% probability that the RMS wavefront is less than 0.105. Figure 2.13 shows the spot diagram and the MTF for sample close to the 90% and 50% performance. Note the system close to an RMS wavefront of 0.065 is nearly diffraction limited. The RMS spot size is 0.905, which slightly larger than the Airy radius. However, there is only a 50% probability that a system will perform better than this if δ = 5 μm. Thus, there's a significant probability that the assembled optical system will perform worse than this.

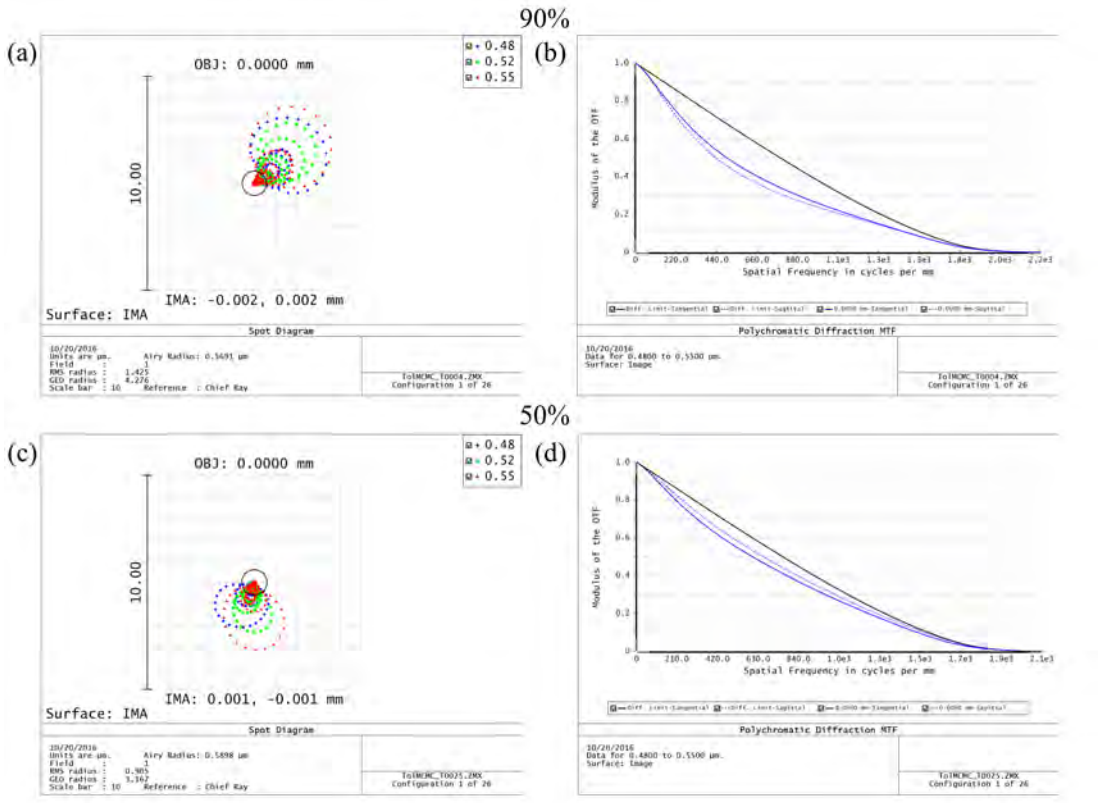

Fig. 2.13 (a) Spot diagram and (b) MTF for sample close to the 90% performance when δ = 5 μm. (c) Spot diagram and (d) MTF for sample close to the 50% performance.

2.7.3 δ = 2.5 μm

The tolerances were reduced to δ = 2.5 μm and analyzed. Once again, 1000 samples are analyzed in the Monte Carlo simulation. For the 1,000 samples the best RMS wavefront is 0.022 and the mean RMS wavefront is 0.038. Based off the Monte Carlo analysis there's a 50% probability that the RMS wavefront is less than 0.037, and a 90% probability that the RMS wavefront is less than 0.052. Figure 2.14 shows the spot diagram and the MTF for a sample close to the 90% and 50% performance. The 50% performance has an RMS spot size slightly less than the Airy radius, and the MTF is nearly diffraction limited, thus there's a 50% probability that the assembled system has near diffraction limited performance. There's a 90% probability that the performance performs better than the system shown in Fig. 2.14(a) and 2.14(b).

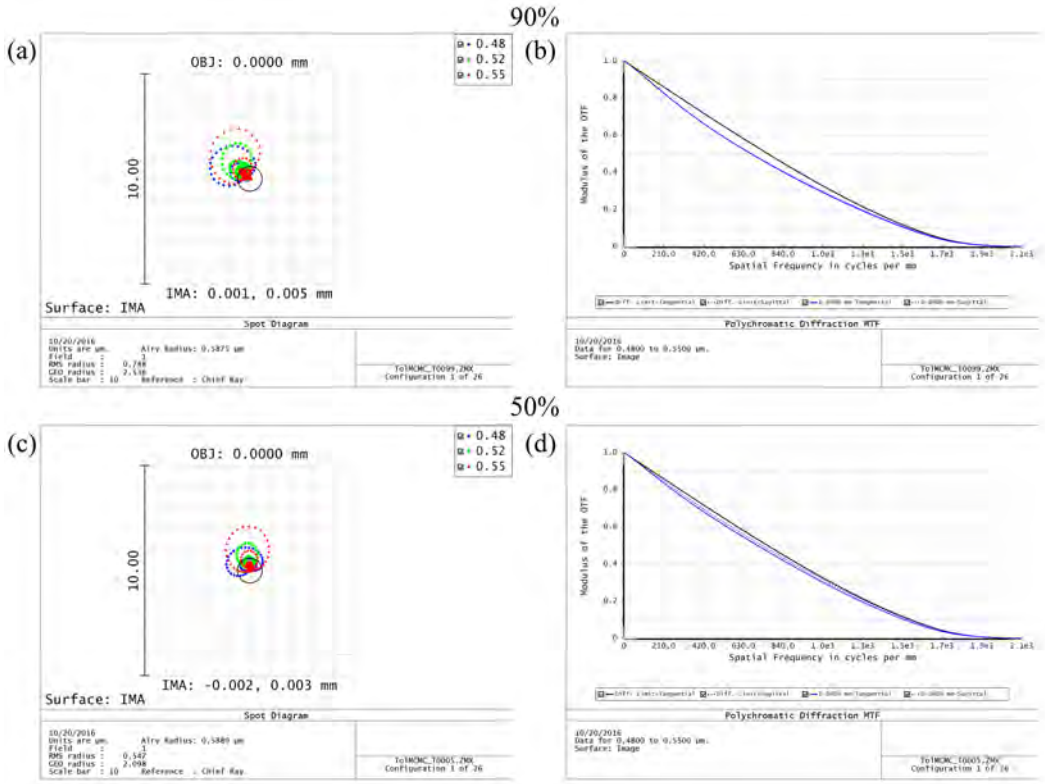

Fig. 2.14 (a) Spot diagram and (b) MTF for sample close to the 90% performance when δ = 2.5 μm. (c) Spot diagram and (d) MTF for sample close to the 50% performance.

Therefore, to guarantee near diffraction-limited performance with a 90% probability, then the decentering between surfaces and elements should be less than 2.5 μm. While not impossible, this is a very tight tolerance.

The situation is complicated even further, since this is an array of microscopes, therefore there's a rotational tolerance between lens elements. If all elements have to be aligned within δ=2.5 μm, that puts a limit on the rotational alignment. If an element a distance *r* away from the center of the rotational axis, an angle of

$$\theta \le \delta / r \tag{2.1}$$

is required to maintain a separation between lens elements less than δ. For an array with lens diameters of 0.5 mm, the longest distance between elements is 0.71 mm. For a maximum separation of δ = 2.5 μm between the centers of the two elements, the angular separation has to be less than 3.52 milliradians.

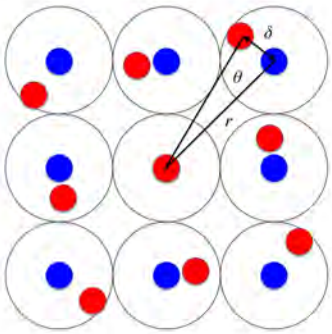

Fig. 2.15 Error caused by rotational misalignment

This section only considered the performance associated with the decentering of lens elements and surfaces. However, this is not the only error that can occur during the fabrication process. In addition to decenters between elements, there is also tilt between elements, errors in the fabricated shape of the surfaces, and errors in the indices of refraction of the elements to name a few. All of these errors will degrade the performance

of the system. This system is very sensitive to external fabrication errors, so the system is not fabricated.

2.8 Conclusion

An array confocal microscope using an array of miniature objectives is designed to determine the feasibility of designing a handheld array confocal microscope. The design of the array confocal microscope consists of a 3×3 array of miniature objectives with a diameter of 0.5 mm. Each objective is designed to focus light on-axis with a 0.5 NA. Each miniature objective array scans a sub-FOV and when all 9 images are combined, the total FOV has an area 9 times the area of the sub-FOV. Therefore, this system is capable of achieving both a high NA and a large FOV. The excitation and detection system is built and works as designed, however, there are significant disadvantages in the design of the scanning system and the objective lenses.

Since this system is designed to be hand-held, the scanning mechanism should be compact, fast and accurate. Unfortunately, a scanning mechanism that achieves all of these requirements doesn’t exist. A translation stage that is a good trade-off between these requirements is identified, and after generous assumptions the maximum driving frequency is 500 Hz. For a 1000×1000 pixel image, this would require 2 s to scan. This is a lower bound on the time required to scan that area, in actuality it would be slower due to finite acceleration of the stages. Thus, it is unlikely that this system can be used to scan objects that have the possibility of moving over time.

Another disadvantage of the system is the tolerances required for the array of miniature objectives to achieve near diffraction-limited performance. By just analyzing the tolerancing required for the decentering of lens elements and surfaces, it is found that

each surface should be aligned within 2.5 μm to achieve near-diffraction limited performance. However, this analysis does not consider the misalignment of the fiber with respect to the objective system, the possible errors in the curvature of the lenses, or tilts between elements. Due to the slow scanning speed and the tight tolerances required for the array of miniature objectives, this system is not fabricated. Instead a different design for an array confocal microscope without these disadvantages is presented in Chapter 4.

3. SINGLE LAYER ONE-DIMENSIONAL BEAM-SPLITTING PHASE GRATINGS

3.1 Introduction

This chapter explains the theory behind the design of continuous phase gratings in the presence of fabrication errors. Continuous phase gratings are ideal components for array confocal microscopes due to their ability to split an incident beam into an array of equally spaced, equal energy beams with high efficiency. This chapter describes the challenges associated with fabricating continuous phase diffraction gratings and describes a design technique that minimizes sensitivity of continuous phase gratings to fabrication errors. With this technique, continuous phase gratings are fabricated with good performance, even in the presence of significant fabrication errors. This chapter is based off one of my publications in Optics Express [53]. The key results from this publication are summarized in this chapter.

There are several methods for splitting one laser beam into multiple beams with equal energy. For example, conventional beamsplitters, diffractive optics, and pupil division have been shown to accomplish amplitude beam-splitting of a laser [54]. Binary phase gratings [55-57], continuous phase gratings [57-59], and refractive lenslet arrays [60] accomplish beam-splitting using phase optical elements. Continuous phase gratings are of interest, since they yield the highest efficiencies out of all beam-splitting methods [61]. As shown in the previous chapter, beam-splitting phase gratings can be used to efficiently split a laser beam into N beams for array confocal imaging. Beam-splitting optics are also used for laser machining and material processing in parallel, sensor systems,

interferometry, communication systems, and image processing and gathering systems [62].

Using calculus of variations, Romero and Dickey found analytical expressions of the optimal continuous phase functions of one- and two-dimensional gratings that maximize energy into N outgoing beams, where N is the desired number of output beams [63,64]. While these gratings theoretically provide the optimal efficiency into N desired output beams, fabrication errors can significantly degrade their performance. For example, Y. Miklyaev *et al.* demonstrated fabrication of the grating profiles described by Romero and Dickey, but fabrication errors prevented equal energy outgoing beams [65].

Instead of optimizing a beam-splitting optical element for maximum efficiency into N outgoing beams, a beam-splitting optical element can instead be optimized for minimum sensitivity to fabrication errors, since fabrication errors degrade the performance of any beam-splitting method. In the presence of fabrication errors, multiple output beams typically no longer have designed energy ratios. For gratings, the grating profile determines the amount by which the energies vary [66]. A grating design with a high sensitivity to fabrication errors may result in an unacceptable variation of the outgoing beam energy ratios. If the sensitivity of the design is low enough, then a lower precision fabrication method could be chosen over a costly, high precision fabrication method.

The continuous phase gratings described in this chapter are fabricated using the maskless photolithography tool (MLT) at University of Arizona [67]. During the MLT fabrication process, two major fabrication errors are identified that significantly degrade the performance of the beam-splitting gratings: errors in the fabricated height and a convolution of the input grating profile with the point spread function (PSF) of the MLT.

3.2 Theory

3.2.1 Fabricated Error in Height

Assume a lossless, one-dimensional grating is characterized by the periodic height function $h(x)$. In the absence of fabrication errors and using a thin grating approximation, the transmitted phase in air of a beam at position x is changed by an amount

$$\phi(x) = h(x)2\pi\left(n-1\right)/\lambda, \tag{3.1}$$

where n is the grating material refractive index at wavelength λ. If there are multiplicative fabrication errors Δ, the accumulated phase has a total phase transmission $(1+\Delta)\phi(x)$. Δ is assumed to be a constant, although in principle any functional form of $\Delta(x)$ can be used. For a grating illuminated by a unit-amplitude plane wave, the output beams, ignoring Fresnel losses, are characterized by the grating's Fourier coefficients:

$$a_p = \frac{1}{T}\int_{-T/2}^{T/2} e^{i(1+\Delta)\phi(x)} e^{-i2\pi px/T}\,\mathrm{d}x \text{ for p = 0, } \pm 1, \pm 2, \ldots \tag{3.2}$$

where T is the period of the grating. The Fourier coefficients determine energies and phases of the output beams [68].

The efficiency of each beam is defined as $|a_p|^2$. A $1\times N$ vector $\boldsymbol{\eta}[\Delta,\phi(x)]$ is defined that contains the efficiency of each desired output beam, where

$$\boldsymbol{\eta}_p[\Delta,\phi(x)] = \left|a_p\right|^2. \tag{3.3}$$

The efficiency of each output beam is dependent on the amount of fabrication error and the phase function of the grating. Total efficiency E of the grating is

$$E[\phi(x)] = \sum_{p=1}^{N}\left|a_p\right|^2, \tag{3.4}$$

where N sums over the number of desired output beams. Note that there may be significant energy in orders outside the desired range, so E might be less than unity. For a 1×9 beam-splitting grating, the total efficiency is the sum of $N = 9$ desired output beams.

Fourier coefficients of an error-free uniform beam-splitting grating are subject to the constraint

$$\left|a_p\right|^2 = E / N, \tag{3.5}$$

which ensures that energies in all desired output beams are equal. If there are fabrication errors, output beam energies in the desired range of p are not equal.

3.2.2 Effect of PSF on Grating Profile

Since the point spread function (PSF) of the writing instrument significantly degrades the performance of the beam-splitting gratings, a phase function that accounts for the PSF of the writing instrument is calculated. If a phase function $\phi(x)$ is desired and the PSF of the writing instrument is given by $h_{\text{inst}}(x)$, the phase function taking the PSF of the writing instrument into account is given by

$$\phi(x) = \phi_{in}(x) * h_{\text{inst}}(x), \tag{3.6}$$

where $\phi_{in}(x)$ is the instrument input function needed to obtain the desired grating phase function. The necessary instrument input function is found by a deconvolution operation. In Fourier-space, the convolution reduces to

$$\Phi(\xi) = \Phi_{in}(\xi) H_{\text{inst}}(\xi). \tag{3.7}$$

$\Phi_{in}(\xi)$ is found by dividing $\Phi(\xi)$ by $H_{\text{inst}}(\xi)$. However, if $H_{\text{inst}}(\xi)$ contains zeroes, there is significant noise amplification during the division. A rect function is multiplied in Fourier-space to act as a low-pass filter and prevent division by zero in the Fourier domain. Since the low-pass filter cuts off higher spatial frequencies, a perfect

reconstruction of the desired phase function cannot be computed. The low-pass filtered instrument input function is given by

$$\Phi_{in}(\xi) = \left(\frac{\Phi(\xi)}{H_{inst}(\xi)}\right) \mathrm{rect}\left(\frac{\xi}{\Delta\xi}\right), \tag{3.8}$$

where $\Delta\xi$ is the width of the low-pass filter. The input phase function in the spatial domain is given by the inverse Fourier transform of $\Phi_{in}(\zeta)$. By rescaling the phase function with a constant *A*, another parameter for optimization is added to compensate for the loss of higher spatial frequencies. The approximate instrument input function that compensates for the PSF is given by

$$\phi_{in}(x) = A\Im^{-1}\left\{\left(\frac{\Phi(\xi)}{H_{\mathrm{inst}}(\xi)}\right) \mathrm{rect}\left(\frac{\xi}{\Delta\xi}\right)\right\}, \tag{3.9}$$

where $\Im^{-1}$ is the inverse Fourier transform and *A* is a scaling factor.

The two main fabrication errors that significantly degrade the performance of beam-splitting diffraction gratings are: (1) fabrication errors in the heights of the gratings Δ and (2) the PSF of the writing instrument $h_{\mathrm{inst}}(x)$. A method to optimize gratings in the presence of fabrication errors in the height and errors caused by PSF of the writing instrument is discussed in the following sections.

3.3 Grating Optimization and Fabrication Results

Since the array confocal microscope described in Chapter 2 requires 9 equal energy beams, 1×9 beam-splitting phase gratings are optimized for good performance over a range of fabrication errors in the height and an error in the estimation of the beam width of the PSF of the writing instrument. The form of the phase function used to design beam-splitting phase gratings into odd numbers of beams is described analytically by [63]:

$$\phi_{design}(x) = \tan^{-1}\left(\frac{Q(x,\boldsymbol{\alpha},\boldsymbol{\mu})}{P(x,\boldsymbol{\alpha},\boldsymbol{\mu})}\right), \tag{3.10}$$

$$P(x,\boldsymbol{\alpha},\boldsymbol{\mu}) = 1 + 2\sum_{j=1}^{M} \mu_j \cos(\alpha_j)\cos(jx), \tag{3.11}$$

$$Q(x,\boldsymbol{\alpha},\boldsymbol{\mu}) = 2\sum_{j=1}^{M} \mu_j \sin\left(\alpha_j\right)\cos\left(jx\right). \tag{3.12}$$

Note that the phase function is parameterized by the phase and amplitude of the corresponding sinusoids.

The reason the PSF has to be considered in the optimization process is shown in the grating profiles in Fig. 3.1. These 3 grating profiles are 1×9 beam-splitting phase gratings that are optimized for a 488 nm laser without considering the PSF of the MTL. This assumes that the PSF is approximated by a delta function, therefore, the gratings are optimized only to minimize the sensitivity to errors in the fabricated height [53]. The blue curves are the designed grating profile and the red dashed lines are the measured fabricated grating profiles using a Bruker Nano NT 9800 optical profiler. In all 3 grating profiles, it is obvious that higher spatial frequencies are attenuated in the fabricated profiles due to the convolution of the PSF with the input grating profile, which results in profiles that are significantly smoothed out. The sharp transitions in the grating profile are not able to be fabricated due to the finite width of the PSF. Furthermore, a measurement of the diffraction pattern for these 3 gratings shows a large deviation from equal energy beams, as shown in Fig. 3.1(b). The grating parameters for these three gratings are shown in Table 3.1, where the peak to valley phase difference for each grating is given by ϕ_{pv}, and E_{design} is the designed efficiency for each grating. From these results, it is obvious that the PSF needs to be considered during the optimization process.

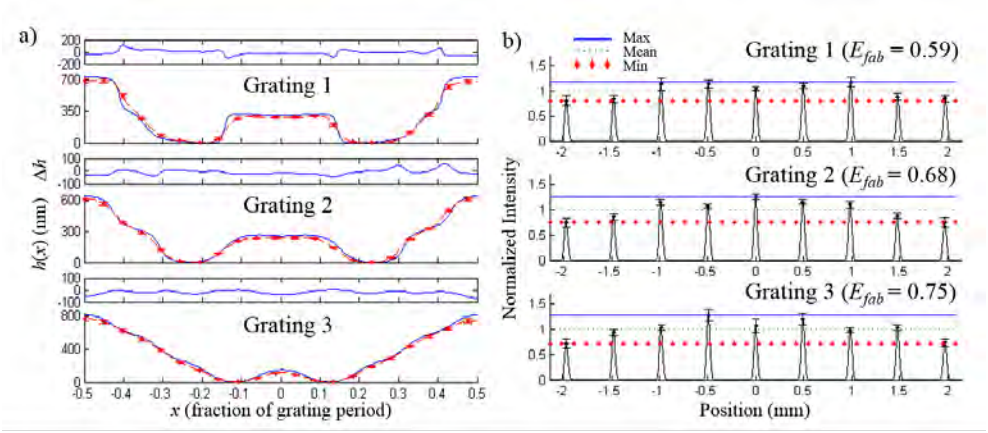

Fig. 3.1 (a) Plot of the design grating profile (solid blue line) and measured fabricated grating profile (dashed red line) in nanometers. Above each grating profile is the difference between the two profiles Δh in nanometers. (b) Measured diffraction pattern for each fabricated diffraction grating is scaled so the mean peak irradiance is equal to unity. For comparison, the red dotted, green dashed, and blue solid horizontal lines are the minimum, mean, and maximum peak irradiances for each grating.

Table 3.1 Grating Specifications

	$\boldsymbol{\alpha}$	$\boldsymbol{\mu}$	ϕ_{pv} (rad)	E_{design}
Grating 1	(4.57, 1.91, 4.61, 5.33)	(1.39, 1.53, 1.14, 0.75)	5.86	0.812
Grating 2	(5.78, 3.56, 4.58, 1.39)	(1.03, 1.36, 1.23, 1.57)	5.12	0.945
Grating 3	(0.72, 5.57, 3.03, 1.41)	(0.971, 0.963, 0.943,1.03)	6.47	0.993

3.3.1 Optimization

The steps for optimizing the design of a phase grating that splits a beam into N equal energy beams with less sensitivity to the PSF of the writing instrument and Δ are as follows: Define a phase function $\phi(x)$ by the parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\mu}$. Deconvolve $\phi(x)$ to obtain $\phi_{in}(x)$. Find a function $\phi_{in}(x)$ when convolved with the estimated PSF, $e^{-\pi x^2/w_{est}^2}$, where w_{est} is the estimated beam width, yields N equal energy outgoing beams, assuming there is no height fabrication error (Δ=0). Since there may be an error in the estimation of the beam width w, the grating is optimized for the w_{est} variable. There are also fabrication errors in the height of the gratings. Therefore, the grating is also optimized to minimize the sensitivity to errors in the height Δ. Variance of the outgoing beam efficiencies is integrated over a range of beam widths and fabrication errors in height. Find a phase function that, when deconvolved, minimizes the integrated variance for the specified

range of fabrication errors. The problem reduces to finding a function $\phi(x)$ that minimizes the cost function

$$C[\phi(x,\boldsymbol{\alpha},\boldsymbol{\mu},A)] = \int_{w_1}^{w_2}\int_{\Delta_1}^{\Delta_2} \sigma^2\{\boldsymbol{\eta}[\Delta,\phi_{in}(x,\boldsymbol{\alpha},\boldsymbol{\mu},A) * e^{-\pi x^2/w^2}]\}\mathrm{d}\Delta\mathrm{d}w \tag{3.13}$$

where σ^2 is the variance of the vector $\boldsymbol{\eta}[\Delta,\phi(x)]$ and $\phi(x)$ is subject to the constraint $|a_p|^2=E/N$ when $\Delta=0$ and $w = w_{est}$. This constraint ensures equal energy output beams assuming the PSF was estimated correctly and there are no fabrication errors in the height. A workflow diagram for the algorithm used to minimize the cost function is shown in Fig. 3.2.

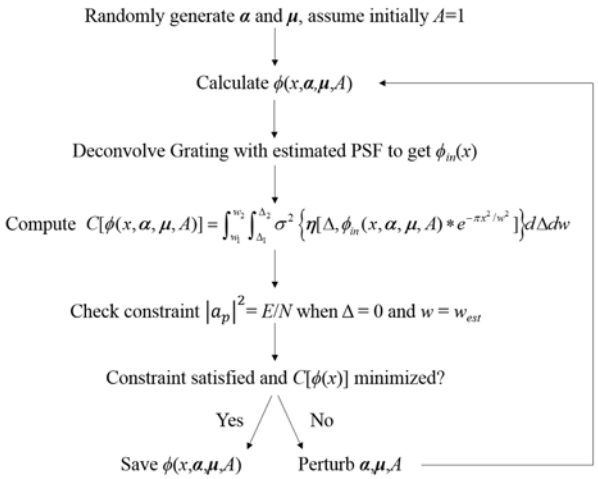

Fig. 3.2 Workflow diagram for optimization algorithm

For the optimization, initial phase functions are generated by randomly choosing $\boldsymbol{\alpha}$ and $\boldsymbol{\mu}$ over [0, 2π] and [0, 4], respectively. The scaling factor A is initially set to unity. The estimated beam width is determined by minimizing the difference between the measured grating profiles from previous fabrication results and a simulated profile when the designed profile is convolved with $\mathcal{h}_{\mathrm{inst}}(x)$. The beam width that minimizes this difference is $w_{est} = 0.0205\,T$, where T is the period of the grating. The phase function is optimized over $\Delta=[-0.1,0.1]$, and the range of beam widths considered in the optimization is $w =$

[0.015T, 0.026T]. The range of beam widths is chosen based off knowledge of the expected fluctuations of the PSF of the writing instrument. The cutoff frequency for the low-pass filter used in the deconvolution is half the Nyquist frequency. Using a constrained nonlinear optimization solver in MATLAB [69], phase functions that satisfy the necessary constraints are calculated. The solver minimizes the cost function while satisfying the equal energy constraint when $\Delta = 0$ and $w = w_{est}$. In order to sample a large enough portion of the parameter space and prevent a search of only local minima, this program is repeated with new initial conditions until 1,000 valid designs are calculated.

Two phase functions to test fabrication are chosen from the 1,000 calculated designs found by the custom optimization program. The specifications for these gratings are listed in Table 3.2. The first design, grating 1, is the phase function with the minimum value for the integrated variance over the fabrication error range. This design has a relatively low $\Delta = 0$, $w = w_{est}$ grating efficiency of E=0.741. Although the entire parameter space is not explored, this phase function should be sufficiently resistant to fabrication errors. The second design, grating 2, has a slightly worse value for the integrated variance, but the $\Delta = 0$, $w = w_{est}$ grating efficiency is E=0.945. This design exhibited the best tradeoff between grating efficiency and insensitivity to fabrication errors. Grating 3 is the optimal efficiency design, designed by Romero and Dickey [63], re-optimized to account for the effect of $\eta_{\text{inst}}(x)$ and has a $\Delta = 0$, $w = w_{est}$ grating efficiency of E=0.993.

Since the optimization algorithm is optimizing over two parameters, a plot of the variance of the output beams is three-dimensional. It depends on the beam width of the PSF and Δ. A three-dimensional plot of the variance of the output beams is displayed in

Fig. 3.3(a). Grating 3 is significantly more sensitive to fabrication errors than grating 1 or grating 2. A density plot of the variance of the output beams is shown in Fig. 3.3(b). The integrated variances for gratings 1-3 are $C[\phi_{design}(x)]\times 10^6$ = 15.3, 47.0 and 157, respectively. Grating 3 is approximately ten times and three times more sensitive than gratings 1 and 2, respectively. The designs are not as sensitive to w as compared to Δ over the ranges chosen.

Table 3.2 Optimized Grating Parameters for Estimated Beam Width w_{est} = 0.0205 T

	$\boldsymbol{\alpha}$	$\boldsymbol{\mu}$	ϕ_{pv} (rad)	E_{design}	$C[\phi(x)]\times 10^6$
Grating 1	(5.31, 4.48, 4.99, 1.16)	(1.16, 1.24, 0.88, 0.55)	6.77	0.741	15.3
Grating 2	(5.78, 3.56, 4.58, 1.40)	(1.03, 1.35, 1.23, 1.57)	5.09	0.945	47.0
Grating 3	(0.72, 5.56, 3.03, 1.41)	(0.973, 0.963, 0.945, 1.03)	6.46	0.993	157

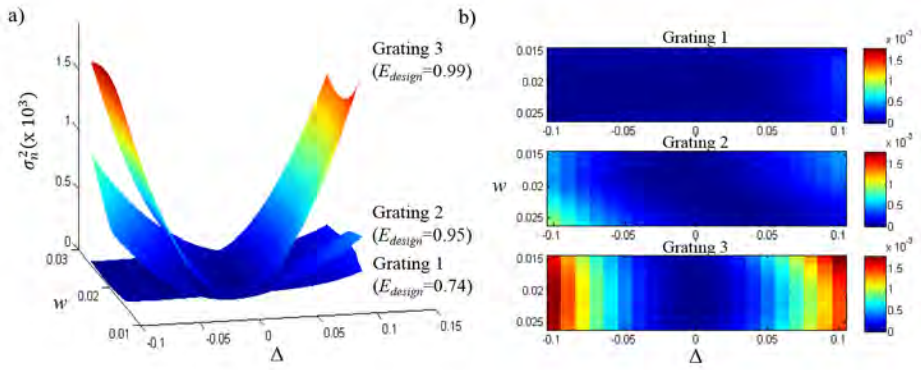

Fig. 3.3 (a) Surface plot and (b) density plot of the variance of the outgoing beams for all 3 gratings as a function of the beam width of the PSF and the fabrication error in the height.

3.3.2 Fabrication and Profile Measurement

All three gratings are fabricated using maskless lithography. The grating period for each of the gratings is 147 μm, and the pixel spacing in the gray-scale photolithographic process is 2.1 μm. The gratings are designed for a wavelength of 488 nm. Height profiles for each fabricated grating are measured using a Bruker Nano NT 9800 optical profiler. The designed height profiles after convolution with the estimated PSF and the average measured height profiles with standard deviations measured at 20 different positions are shown in Fig. 3.4(a). The average peak-to-valley heights for each fabricated grating are 780 ± 24 nm, 650 ± 12 nm, and 807 ± 23 nm. The designed heights for each grating are

854 nm, 643 nm, and 815 nm, respectively. The fabrication error for each fabricated grating is calculated from the height measurements using

$$\Delta_{fab} = \frac{h_{fab} - h_{design}}{h_{design}}, \tag{3.14}$$

and is listed in Table 3.3. The difference between the average measured height profiles and the designed height profile Δh is also shown in Fig. 3.4(a). From the measured grating plots, the ringing in the grating profiles is indicative that the beam width of the PSF was estimated incorrectly. By analyzing the measured grating profiles, the beam width during the writing process was determined to be $w_{fab} = 0.017T$, instead of the assumed value $w_{est} = 0.0205\ T$ used during the optimization. The incorrect beam width estimation results in the ringing seen in gratings 1 and 2. Also, note the significant errors in the fabrication of grating 2, where the fabricated profile is steeper than the theoretical profile. Slices of the plots in Fig. 3.5 at the fabricated beam width $w_{fab} = 0.017T$ are shown in Fig. 3.5.

3.3.3 Testing

Each grating is tested by illuminating with a collimated laser beam and placing a detector at the focal plane of a 150 mm lens. A profile of the diffraction pattern is measured using a BladeCam-XHR beam profiling camera. The fabricated grating is about 100 times larger than the incident laser beam, so diffraction patterns are measured when the incident beam is placed at 20 different locations on the grating. The diffraction patterns for each grating are shown in Fig. 3.4(b). The average measured diffraction patterns are normalized by scaling the mean peak irradiances to unity. Error bars on the peak irradiance of each outgoing beam are the standard deviations of the peak irradiances for all 20 measurements. After normalization, the standard deviations of the measured peak

irradiances for each of the 20 measurements are calculated. The standard deviations σ_{fab} are 0.088 ± 0.018, 0.058 ± 0.009, and 0.140 ± 0.064 for gratings 1-3, respectively. To compare measured performance of the fabricated gratings to theoretical performance, variance is calculated after scaling the mean peak irradiances to the designed E/N. The measured variances after scaling to the designed E/N and range of fabrication errors for each grating are indicated in Fig. 3.5. Note that the diffraction pattern for grating 2 is very close to equal energy, as shown in the zoomed-in portion of the variance of grating 2 in Fig. 3.5. The differences δ_{fab} between the minimum and maximum scaled peak irradiances is calculated for each of the 20 locations and are 0.252 ± 0.042, 0.168 ± 0.030, and 0.396 ± 0.170 for gratings 1-3, respectively.

The absolute diffraction efficiency of each grating is measured by calculating the ratio of energy in the desired outgoing beams to the energy in the incident laser beam. The measured absolute diffraction efficiencies for gratings 1-3 are 0.628 ± 0.020, 0.729 ± 0.049, and 0.729 ± 0.013, respectively. In the presence of fabrication errors, grating 3 has the same measured diffraction efficiency as grating 2. The measured absolute diffraction efficiency is reported in Table 3.3 as E_{fab}, which is lower than the design diffraction efficiency for several reasons. The first reason for a reduced E_{fab} is due to Fresnel losses and absorption. A measurement resulted in 18 ± 2% of the incident light being lost due to Fresnel losses and absorption. Also, in the presence of fabrication errors in the grating profile, energy is transferred from the desired 9 outgoing beams into beams outside the desired range. A third reason is due to tool signature in the direction perpendicular to the grating profile, which causes slight fluctuations in the height of the grating in the perpendicular direction. Due to the tool signature, light diffracts in a direction

perpendicular to the grating profile. This diffraction can result in multiple low amplitude copies of the diffraction pattern from the grating in the perpendicular direction, reducing the diffraction efficiency into the desired 9 outgoing beams.

Table 3.3 Grating Performance for w_{est} = 0.0205 T

	E_{design}	E_{fab}	h_{design} (nm)	h_{fab} (nm)	Δ_{fab}	σ_{fab}	δ_{fab}
Grating 1	0.741	0.628 ± 0.020	854	780 ± 24	−0.086 ± 0.012	0.088 ± 0.018	0.252 ± 0.042
Grating 2	0.945	0.729 ± 0.049	643	650 ± 12	0.010 ± 0.003	0.058 ± 0.009	0.168 ± 0.030
Grating 3	0.993	0.729 ± 0.013	815	807 ± 23	−0.010 ± 0.019	0.140 ± 0.064	0.396 ± 0.170

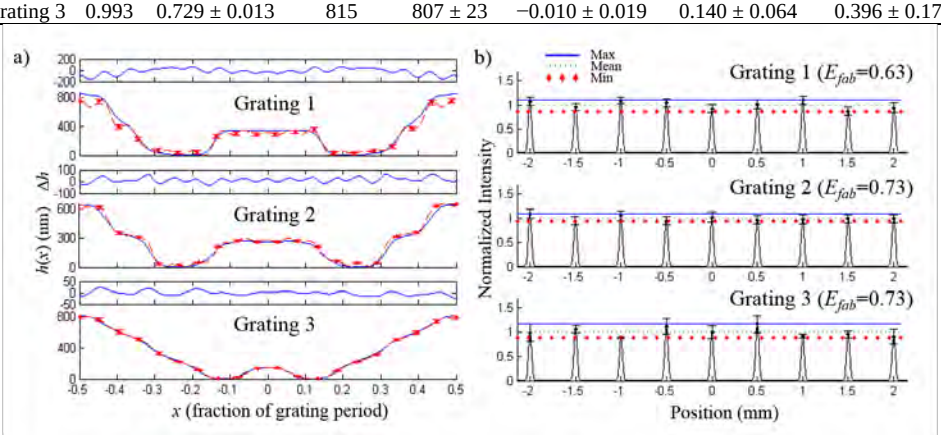

Fig. 3.4 (a) Plot of the design grating profile after convolution with the estimated PSF (solid blue line) and measured average fabricated grating profile (dashed red line). Above each grating profile is the difference between the two profiles Δh in nanometers. (b) Measured diffraction pattern for each fabricated diffraction grating is scaled so the mean peak irradiance is equal to the designed beam efficiency. For comparison, the red dotted, green dashed, and blue solid horizontal lines are the minimum, mean, and maximum peak irradiances for each grating.

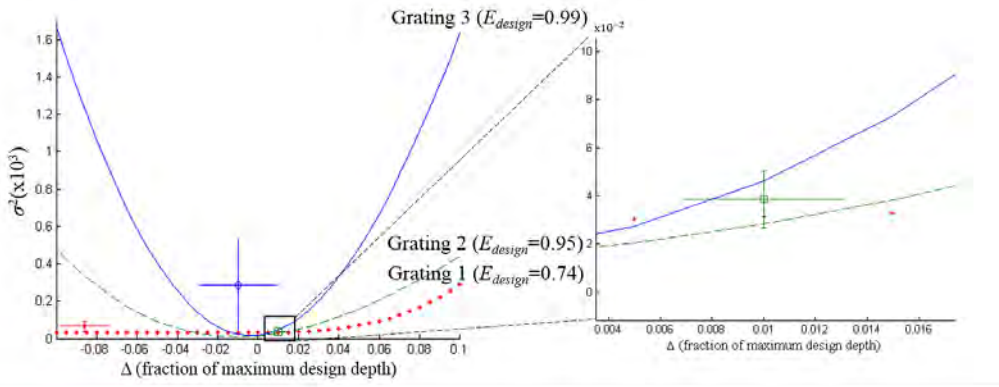

Fig. 3.5 Sensitivity plot for the three theoretical convolved gratings with estimated beam width of 0.017T; grating 1 (red dotted line), grating 2 (green dashed line), and grating 3 (blue solid line). The variance of the energy of the outgoing beams is plotted as a function of the fabrication error Δ as a beam travels through a theoretical convolved grating. Horizontal and vertical error bars correspond to $\sigma^2\{\eta[\Delta_{fab}, \phi_{fab}]\}$ for the measured range of Δ_{fab}. The box on the left graph is expanded on the right graph, where the measured variance and range of Δ is shown for grating 2.

By taking $\eta_{\text{inst}}(x)$ into account during optimization, the performance of all three gratings is improved, even in the presence of a 20% error in the estimation of the beam

width of the PSF. Grating 2 has the best performance and is the grating that is used in Chapter 2. By looking at the measured profile for grating 3 in Fig. 3.4(a), it can be seen that the average profile is very close to the theoretical profile, but the diffraction pattern is very far from equal energy output beams. If grating 2 is considered, there are significant fabrication errors, notably the ringing in the middle portion of the grating and the incorrect steepness of the edge of the grating; however, the performance of this grating is very close to equal energy output beams. For grating 1, there is a very large error in the height of the grating profile, yet it still performs better than grating 3. Therefore, the optimization algorithm developed in this chapter creates grating designs that are robust to fabrication errors in height and can be used to account for the PSF during the writing process. While grating 3 theoretically provides the optimal efficiency of light into the desired 9 outgoing beams, this chapter demonstrates that this design is extremely sensitive to fabrication errors and may be too difficult to fabricate. Therefore, a design that has a slightly lower efficiency, but is more robust to fabrication errors is more desirable.

3.4 Conclusion

In conclusion, two significant factors that degrade the performance of continuous phase gratings are identified, which are errors in the height of the grating profiles and errors caused by the finite beam width of the PSF during the writing process. A method for optimizing the grating design in the presence of these fabrication errors is developed. These optimized grating designs sacrifice total efficiency, but they provide lower sensitivity to fabrication errors. Using the numerical results for a 1×9 beam-splitting

grating, three different gratings are fabricated. Two gratings are optimized for fabrication and are compared to the grating optimized for maximum efficiency.

An optimization algorithm that reduces sensitivity to fabrication errors in the height and sensitivity to the PSF of the writing instrument is developed. Three gratings are fabricated using the results from this optimization algorithm. Both gratings using the developed algorithm perform better than the optimal efficiency grating after fabrication. Measured performance of the best grating fabricated using the developed algorithm is 3.4 times better than the optimal efficiency grating, with only a slight reduction (0.945 compared to 0.993) in the theoretical efficiency. The measurement of σ_{fab} is 0.058 ± 0.009 for the best grating fabricated, whereas σ_{fab} is 0.201 ± 0.047 for the optimal efficiency grating. Therefore, in the presence of fabrication errors, the optimal efficiency design may not be desirable since the high sensitivity to fabrication errors may yield too large of a deviation from equal energy output beams. Depending on the accuracy of the writing instrument, a less efficient grating design with less sensitivity to fabrication errors may be desirable. This chapter demonstrates an effective algorithm to design gratings with less sensitivity to fabrication errors.

Instead of using the optimal efficiency grating (grating 3) for the array confocal microscope in Chapter 2, grating 2 is a much better choice since it not only has the same fabricated efficiency as the optimal efficiency grating, but it also has better uniformity between the diffracted beams. This uniformity between diffracted beams is desired since the diffracted beams are coupled to an array of fibers. Ideally, the coupling efficiency is the same through all 9 fibers, however, the coupling efficiency in each fiber decreases if the focused beams are not aligned perfectly with the fiber and/or the incident beams have

unequal energy. However, the optimization algorithm allowed for the successful fabrication of a beam-splitting diffraction grating with significant fabrication error that allowed for relatively high coupling through the fiber array. The coupling efficiency through the fiber array was measured to be 65.5 ± 2.9%, with a maximum energy difference between beams of 15 ± 2%. The maximum energy difference between beams after exiting the fiber array matches δ_{fab} for grating 2, which is 16.8 ± 3.0%. This suggests the beams are well aligned with the fiber array. Therefore, the combination of a beam-splitting continuous phase grating optimized for reduced sensitivity to fabrication errors and a fiber array is an efficient method of splitting *N* beams into a geometry that is useful for array confocal microscopy.

4. WHOLE BRAIN IMAGING WITH AN ARRAY CONFOCAL MICROSCOPE

4.1 Introduction

Reconstructing wiring diagrams in the brain still remains a significant challenge for researchers. Wiring diagrams in the brain are extremely important, since they may reveal structural or molecular circuit differences that differentiate normal versus diseased brains. A fundamental understanding of these differences could potentially lead to developments in the treatment, prevention, or cure for brain disorders like Alzheimer's, schizophrenia, autism, epilepsy, and traumatic brain injury.

One common method to reconstruct wiring diagrams is to slice the brain into sectioned tissue [70-72]. The sectioned tissue has to be thin enough for optical microscopes to image the entire volume without scattering degrading the image quality significantly. Scattering in the tissue is the factor that limits the penetration depth of microscopes. By using a confocal microscope, the optical sectioning capabilities allow good image quality up to a couple hundred microns into the tissue. Therefore, the sections have to be less than a couple hundred microns thick to image using a confocal microscope. While a two-photon microscope penetrates deeper than a confocal microscope, the penetration depth is still limited. For whole brain imaging using two-photon microscopes, the brain still has to be sectioned into layers [72].

Not only is the process of fully sectioning a brain time-consuming, but there is also an inherent disadvantage in studying the brain using this method. During the sectioning,

pathways in the brain may become severed, leading to difficulty in accurately aligning pathways between two adjacent sections. Furthermore, aligning long-range pathways over several millimeters in depth becomes extremely difficult. Therefore, it is extremely challenging to fully understand the 3D structure of the brain using sectioned tissue. While understanding these wiring diagrams at the microscopic level is important, to fully understand the brain the macro and microscopic data needs to be merged for a system level understanding of the function of the brain [73].

In recent years, several brain clarification methods, such as CLARITY [74], CUBIC [75], BABB [76], and Scale [77], have been demonstrated. Brain clarification methods are processes in which the brain's structural and molecular information is retained, but the brain is optically transparent and macro-molecule permeable. Scattering is significantly reduced using the brain clarification techniques; therefore, conventional fluorescence microscopes can penetrate much deeper into the brain.

Even though the brain clarification methods solve the problem of penetration depth for brain imaging, microscopes are typically not suited for whole brain imaging. As the numerical aperture (NA) of the objective increases, the field of view (FOV) and the working distance decrease. Therefore, high NA objectives often have short working distances, which make it impossible to image the entire brain, since the brain will hit the first surface of the objective before imaging the bottom of the brain. For example, the microscope used by Chung et al. to image a 5-6 mm thick mouse brain using single photon excitation had a working distance of 3.6 mm [74]. To image the whole brain, first the dorsal half of the brain was imaged, then the brain was inverted and the ventral half was imaged. Long working distance objectives have been made, however, the FOV is

typically small. The brain would then have to be translated by the FOV of the objective to cover the entire region of the brain.

As an example, suppose an objective has a FOV of 0.5 mm × 0.5 mm, and the lateral dimensions of a rat brain is 11 mm × 11 mm. The brain needs to be translated 484 times assuming no overlap between each successive image to cover the entire lateral area of the brain. Now suppose an objective is designed with a FOV of 3 mm × 3 mm. The objective only needs to be translated 16 times to cover the same region. Therefore, it is extremely advantageous to have a larger FOV to significantly speed up the imaging of the brain. This problem is compounded even further, since at every lateral position on the brain, the brain is also scanned axially. Suppose at every position there are N_z axial movements, then the brain with the smaller FOV has to be translated $484N_z$ times, whereas the larger FOV only has to be translated $16N_z$ times. Therefore, the larger FOV has the potential to significantly reduce the total acquisition time for whole brain imaging.

This chapter details the design and testing of an array confocal microscope built for whole brain imaging. The microscope objective is specifically designed for whole brain imaging with a FOV of 3 mm × 3 mm and a working distance of 15.5 mm. The microscope is designed with a large working distance to image cleared brains, so the whole brain can be imaged without flipping or sectioning. To speed up the acquisition time, an array of 9 beams are used to scan the FOV 3 mm × 3 mm to reduce the scanning time over that FOV by a factor of 9.

4.2 Design

The array confocal microscope is designed to have a 0.5 NA objective, a 3 mm × 3 mm FOV, and 15.5 mm working distance. The excitation wavelength is 488 nm, and the

optical components are optimized for a wavelength range of 480 nm – 560 nm. The layout of the microscope is shown in Fig. 4.1. A summary of the design is as follows: a 488 nm argon-ion laser is split into 9 equal energy beams using beam-splitting gratings. Each of these 9 beams is coupled into a single-mode fiber. A mechanical mount is designed, such that the light from each single-mode fiber is reflected off a dichroic mirror and all 9 beams converge onto a galvo-scanning mirror at specified angles. A relay system relays the beams into the entrance pupil of the objective, and the 9 beams are focused at the sample plane, where the 9 beams are spatially separated in a 3×3 array. The sample is excited by the 9 beams and fluoresces, and the fluorescent light travels back through the objective and relay system, traversing the reverse path of the excitation light. The florescence transmits through a dichroic mirror, where the fluorescent light is focused onto the end of a multi-mode fiber. The light exiting from each multi-mode fiber is relayed onto a PMT array, where the intensity of the fluorescent signal is measured and transferred to a computer. Each component of the microscope is detailed in the following sub-sections in this chapter.

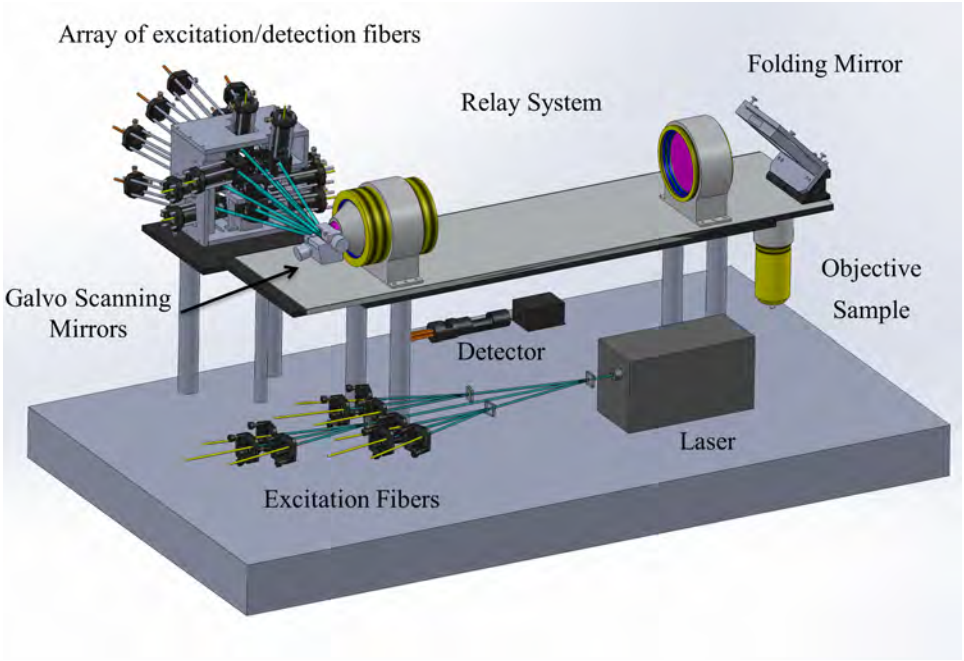

Fig. 4.1 Layout of the array confocal microscope

4.2.1 Coupling to Excitation Fibers

A 1×3 binary Dammann grating is used as the beam-splitting optics for this microscope [62]. The Dammann grating has a phase ϕ_0 = 2.008 radians and the phase profile is shown in Fig. 4.2(a). The height of the grating is calculated using

$$h = \frac{\phi_0}{2\pi} \frac{\lambda}{n(\lambda) - 1} \tag{4.1}$$

where λ is the wavelength, and n is the index of refraction at wavelength λ. For this microscope λ = 488 nm and the Dammann grating is made from fused silica. The index of fused silica is n = 1.4630 at λ = 488 nm. Therefore, the designed height of the grating is 336.8 nm. To determine the period of the grating T, the grating equation is used assuming normal incidence, which is

$$\sin\theta_m = \frac{m\lambda}{T} \tag{4.2}$$

where θ_m is the output diffraction angle for diffraction order m. The grating is designed to have a diffraction angle of 9.5°, so T = 2.943 μm. The designed height profile is shown in Fig. 4.2(b) and the resulting diffraction pattern of the Dammann grating is shown in Fig. 4.2(c). Note the diffraction pattern consists of three equal energy beams diffracted at an angle of 9.5°.

The layout to split the incident laser beam into 9 outgoing beams is shown in Fig. 4.2(d). Collimated light from a laser is incident on a Dammann grating, which splits the beam into 3 equal energy collimated output beams. Each of these three beams travels for some distance and is then incident on another Dammann grating. Therefore, each of these three beams is split into three more beams. An adjustable aspheric fiber collimator is mounted on a tip/tilt stage and the light from each of the 9 beams is coupled into a single-mode fiber. Ignoring absorption losses and Fresnel reflections, theoretically 7.8% of the incident light reaches the single-mode fiber based off of diffraction efficiency. Experimentally, 5% ± 1% of the incident light is transmitted through each fiber.

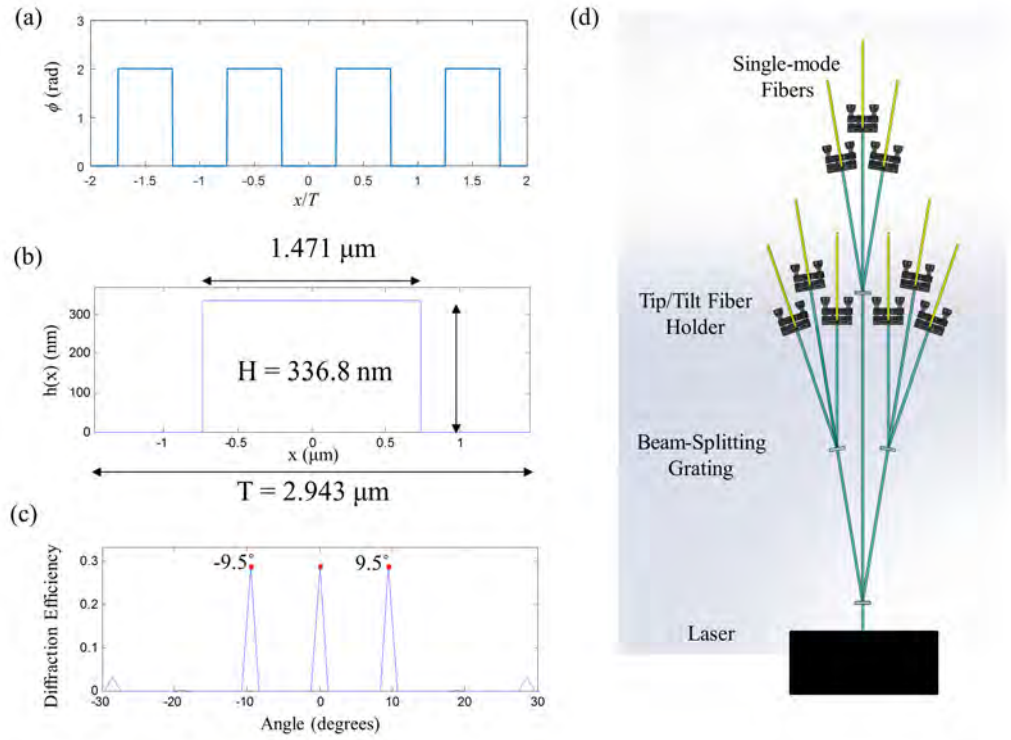

Fig. 4.2 (a) Phase profile for 1×3 Dammann grating. (b) Height profile for fabricated grating. (c) Far-field diffraction pattern. (d) Layout for single-mode fiber coupling for 9 beams.

4.2.2 Mechanical Mount

The other end of each of the 9 single-mode fibers is attached to a confocal arm as shown in Fig. 4.3(a). The excitation fiber acts as the illumination pinhole for the confocal microscope and a multimode fiber acts as the detection pinhole. The light from the single-mode fiber is collimated using a 30 mm lens positioned inside a lens tube and is clipped to a diameter of 7 mm using a mask. The beam reflects off a dichroic mirror and is directed to reflect off a 2-axis galvo scanner.

In order for the beams to be separated by 1 mm in the *x* and *y* direction at the sample plane, the beams need to be separated by an angle of 9° when incident on the galvo. A mechanical mount, as shown in Fig. 4.3(b), angles each of the confocal arms to create the required angular separation between each of the 9 beams.

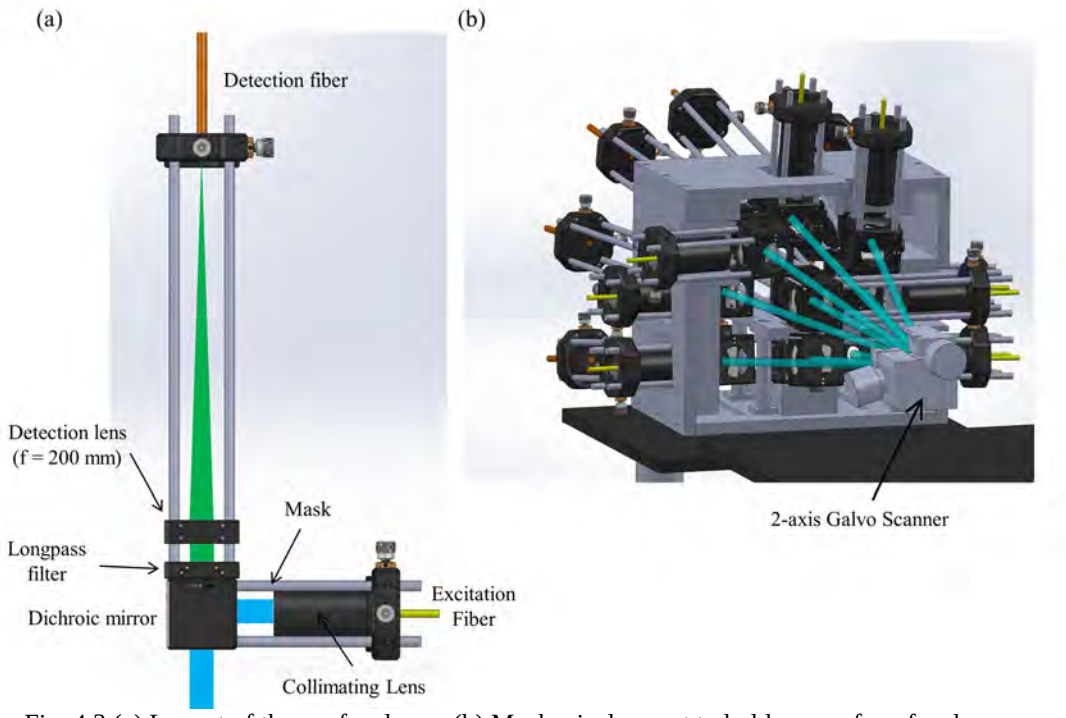

Fig. 4.3 (a) Layout of the confocal arm. (b) Mechanical mount to hold array of confocal arms.

The Solidworks model for the mechanical mount is shown in Fig. 4.4(a) and an exploded view of the model is shown in Fig. 4.4(b). The sides of the mechanical mount are tilted, such that each confocal arm sends the excitation light at the required angle. Square holes are positioned on the mechanical mount to allow the cage rod assembly for the excitation fiber to be positioned correctly. A slight complication arises for the central and bottom middle confocal arms since a lens tube would block the excitation beam if short cage rods are used. Longer cage rods are used for these confocal arms, and the excitation beams from the central right and bottom right confocal arms pass through the cage rods of the central and bottom middle confocal arms.

To compensate for any errors in the fabrication of the mechanical mount, the holes to attach the confocal arms to the mechanical mount are slightly larger than required, so the confocal arm can be rotated slightly with respect to the normal to the plane of the mechanical mount. Additionally, shims can be positioned between the confocal arm and

the mount to tilt the confocal arm about the plane of the mount. These two compensation methods are used to tilt each confocal arm to the designed angle with respect to the galvo scanner. This mechanical mount is 3D printed using a Makerbot 3D printer, and the 3D printed mount is shown in Fig. 4.4(c). Figure 4.4(d) shows the mechanical mount with the dichroic mirrors and parts of the cage system attached, and Fig. 4.4(e) shows the mechanical mount during alignment.

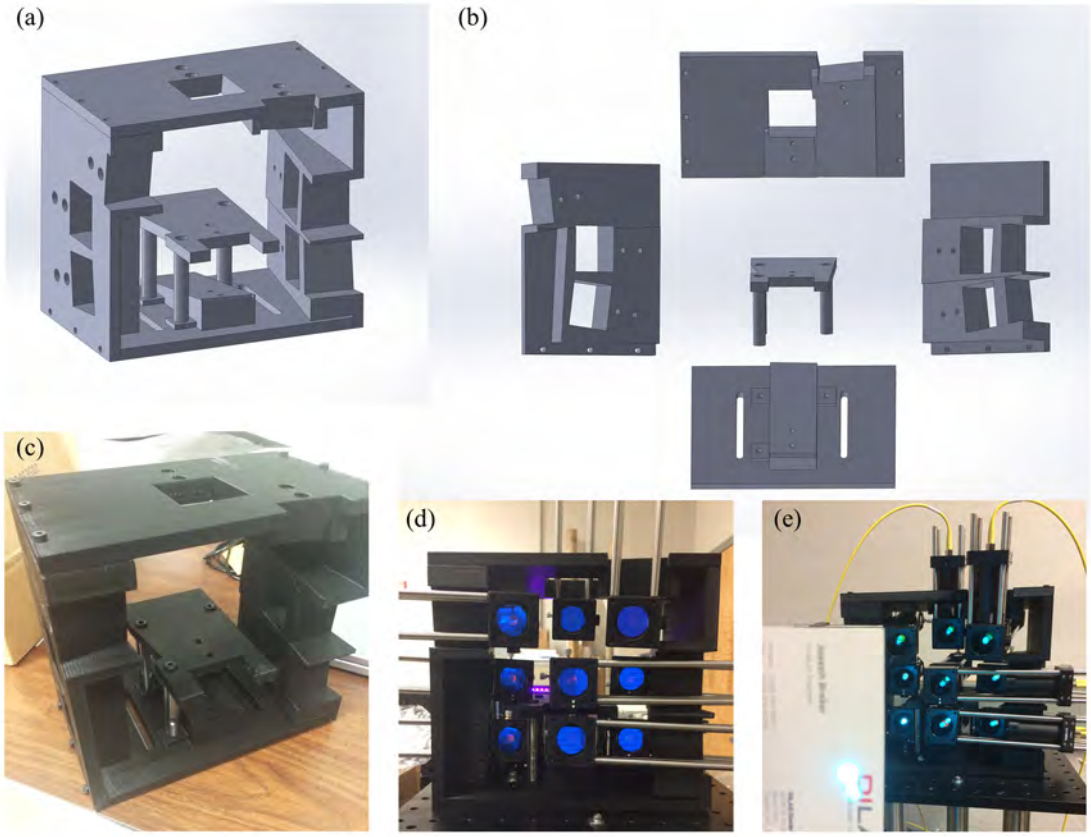

Fig. 4.4 (a) Solidworks model of mechanical mount (b) Exploded view of Solidworks model. (c) Photo of 3D printed mount. (d) Front view of assembled 3D printed mechanical mount with dichroic mirrors. (e) Photo of mechanical mount during alignment.

4.2.3 Relay System and Objective

The optical design of the relay system and objective is shown in Fig. 4.5. The beams are separated by 9° after the galvo system. Through the relay system, the beams are enlarged from a diameter of 7 mm to a diameter of 22 mm after the galvo scanning mirrors. All 9

beams are relayed to the entrance pupil of the objective, where all 9 beams are then focused to a 3×3 array at the sample plane, as shown in Fig. 4.5.

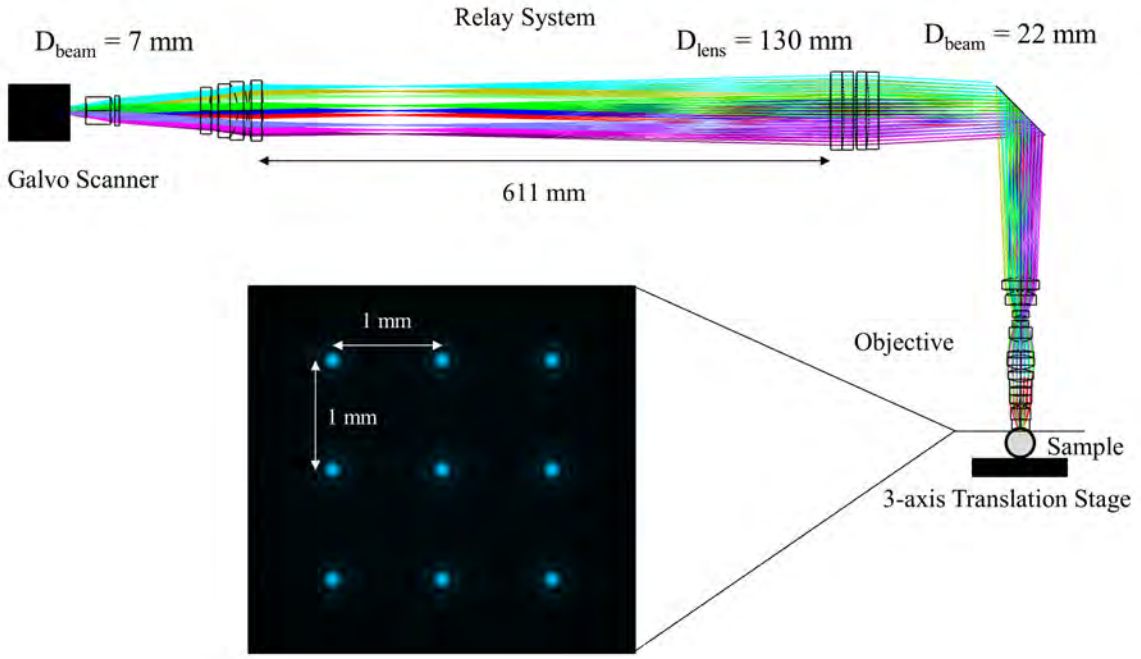

Fig. 4.5 Optical layout for relay system and objective.

The optical design for the microscope objective is shown in Fig. 4.6(a). The objective is designed to have a 0.5 NA and is designed to image a 3 mm × 3 mm FOV with a 15.5 mm working distance. It's an oil immersion objective, designed for oil with an index of n_d = 1.45. The objective is optimized for the wavelengths from 480 nm – 560 nm. As shown in Fig. 4.6(b), the MTF shows near diffraction limited performance over the entire FOV.

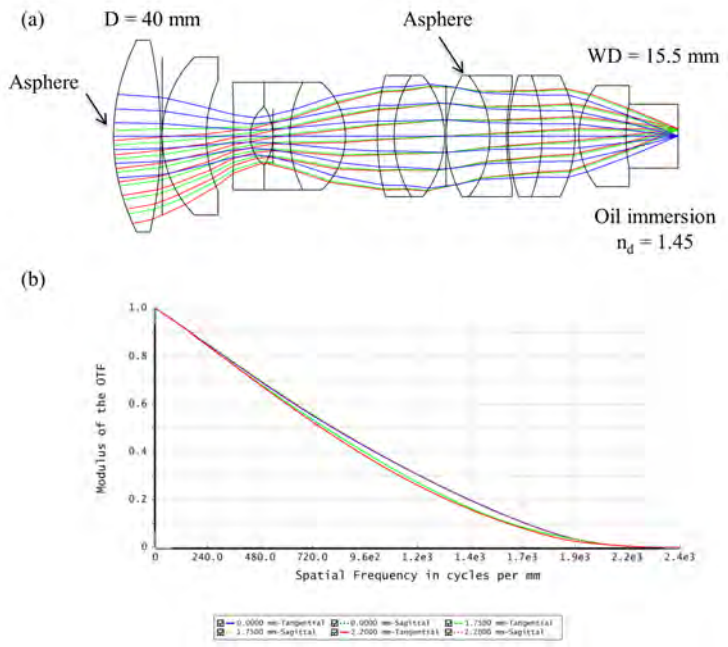

Fig. 4.6 (a) Optical design for microscope objective. (b) MTF for objective, showing near diffraction-limited performance over the entire FOV.

The mechanical mounts for the relay system and the objective are shown in Fig. 4.7, note the large size of the mounts. The relay system has mounts with a diameter of 155 mm and the objective barrel has a diameter of 70 mm. The size of a typical microscope objective is shown next to the microscope objective as a comparison of the size. This system is much larger than a conventional microscope.

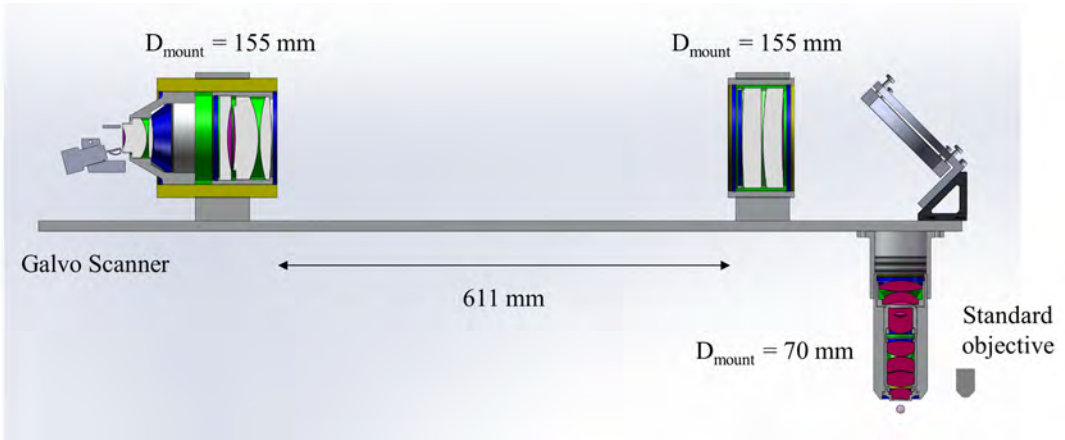

Fig. 4.7 Solidworks model for mechanical mounts for relay system and objective

4.2.5 Detection

The fluorescence from each of the 9 beams travels back through the objective and relay system. After reflecting off the galvo scanning mirror, the fluorescence from each beam returns to its respective confocal arm. The fluorescent light is transmitted through a dichroic mirror and is transmitted through a longpass filter that transmits wavelengths greater than 500 nm. A 200 mm lens is used to focus the fluorescence onto the end of a multimode fiber, which acts as the detection pinhole, as shown in Fig. 4.8(a). Assuming a 12.5 mm diameter aperture, after focusing from 200 mm lens (Thorlabs AC-254-200-A), the polychromatic Airy disk radius is 9.37 μm over the wavelength range 480 nm – 560 nm. The fiber diameter is 25 μm, which is 1.33 times the Airy radius. Since the choice of the diameters of the fiber and focusing lenses are discrete (10 μm, 25 μm, or 105 μm (Thorlabs custom fiber HPSC)) the options for focusing onto the end of the multimode fiber are limited. To focus onto the end of the 10 μm diameter fiber, a 50 mm focusing lens needs to be used, however, the performance of the lens (Thorlabs AC-254-050-A) is not diffraction limited. To closely match the 25 μm diameter, a 275 mm lens needs to be used, but the track length would be longer than the cage rods used. Due to the limited number of fiber diameters and off-the-shelf focusing lenses, a fiber diameter is chosen with slightly larger Airy radius than 1, since an Airy radius smaller than 1 doesn't increase resolution, but it does significantly decrease the light efficiency. Another reason to choose a slightly larger fiber diameter is light efficiency. Theoretically, the light efficiency from the sample plane to just before the dichroic mirrors in each channel is 41% at 530 nm assuming MgF_2 coatings as anti-reflection coatings in Zemax. This calculation doesn't consider the transmission through the dichroic mirror, coupling to the

detection fiber, absorption losses in the multimode fiber, or imaging through the 1:1 detection relay system.

Each multimode fiber is attached to the mount shown in Fig. 4.8(b). This mount holds each of the fibers, such that the centers of the fibers overlap with one channel of the PMT array after a 1:1 relay system, as shown in Fig. 4.8(c). The dimensions of each PMT channel are 0.8 mm × 16 mm. Due to the large width of each PMT channel, the 1:1 relay does not require diffraction-limited performance, as long as the fluorescence from each multimode fiber is within each channel of the PMT array. The aberration that is important to control is distortion, since a misalignment between the multimode fibers and the PMT array causes decreased efficiency in each channel and possible crosstalk. The optical diagram for the 1:1 relay is shown in Fig. 4.9(a), the spot diagram for this system is shown in Fig. 4.9(b), and the footprint diagram is shown in Fig. 4.9(c). Note that even though the spot size is larger than the diffraction-limited spot size, it is still within the dimensions of a channel in the PMT array.

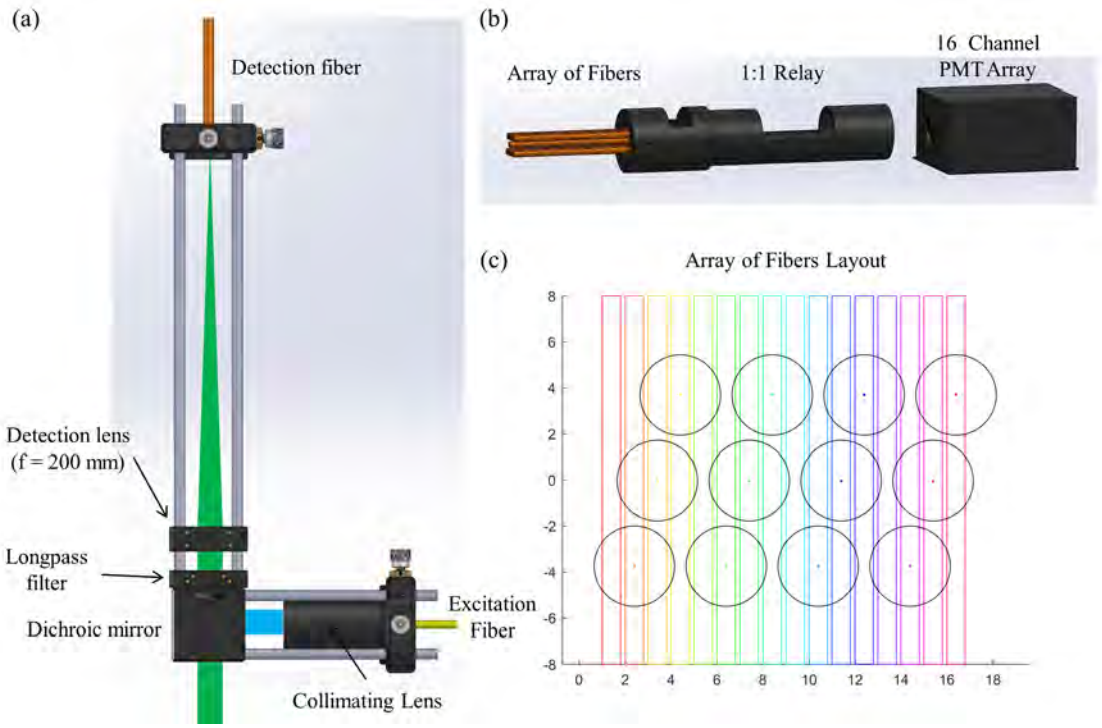

Fig. 4.8 (a) Fluorescence path in confocal arm. (b) Mount for array of fibers, 1:1 relay, and PMT array. (c) Design for relay system where the centers of each fiber are centered on a single channel in the PMT array.

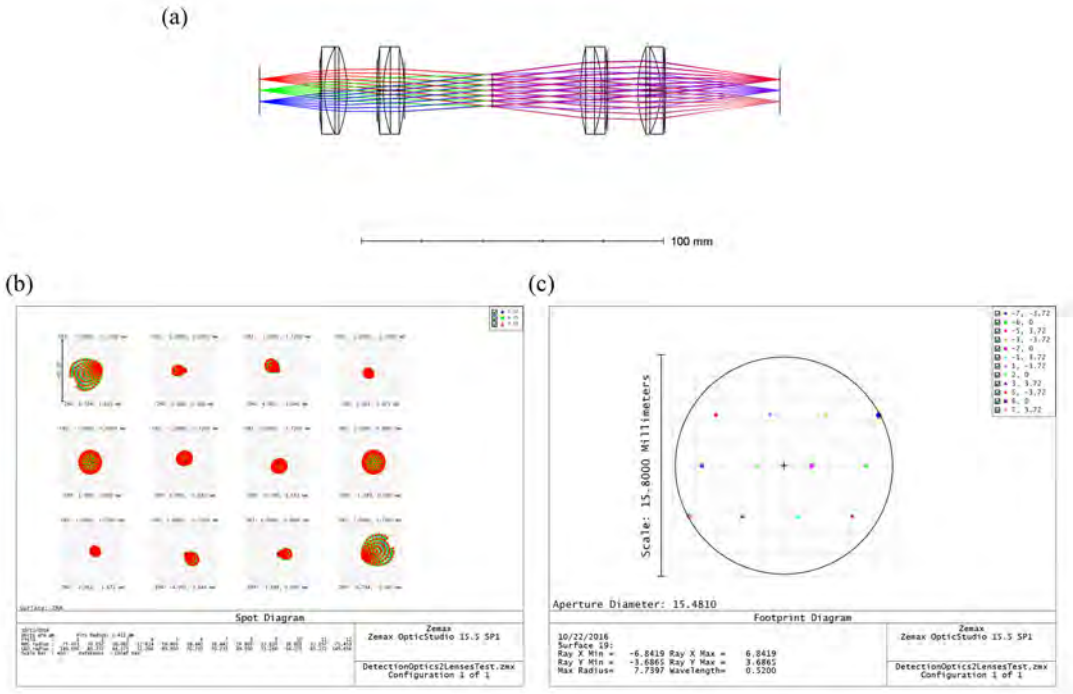

Fig. 4.9 (a) Optical layout for the 1:1 relay used in the detection system. (b) Spot diagram and (c) footprint diagram for relay system.

4.2.6 Image Acquisition Software

A custom Labview program is developed in order to control the scanning of the galvo mirrors, the movement of the translation stage, and the signal readout from the PMT array. The user interface for the Labview software is shown in Fig. 4.10. The inputs into the software are the number of rows N_r and columns N_c for each channel in the microscope. The image size generated is $3N_r \times 3N_c$. Additionally, the user inputs the requested dwell time for each pixel. However, the actual pixel dwell time may be determined by the maximum frequency of the galvo scanner. If a pixel dwell time is smaller than is allowed based off the maximum speed of the galvo, then the actual pixel dwell time is slower than requested.

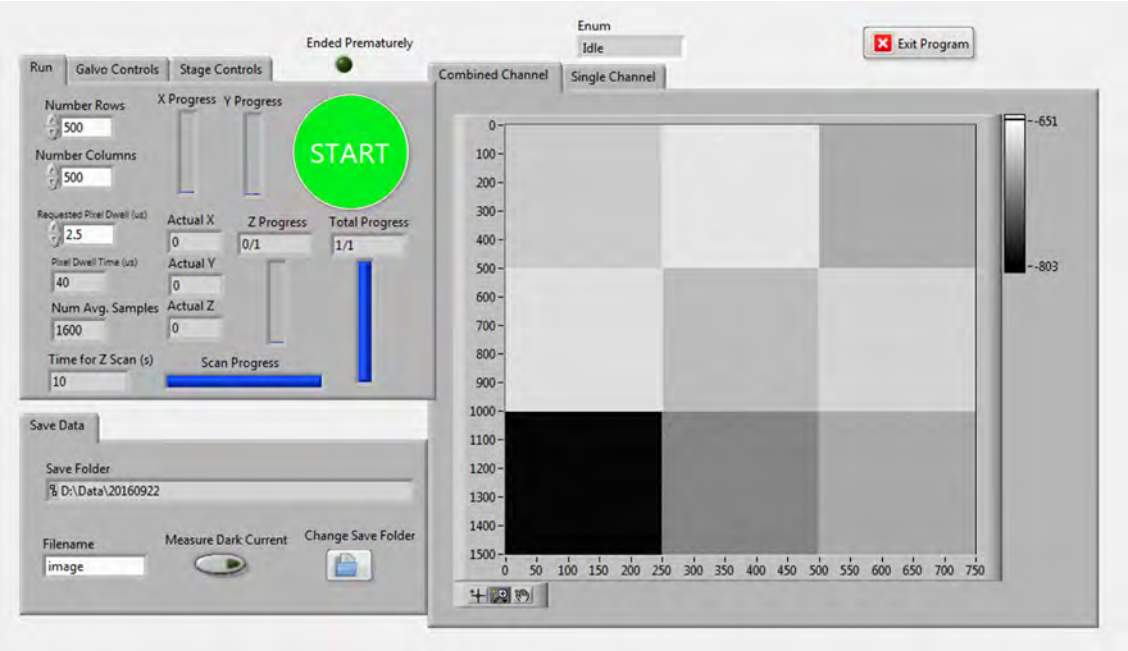

Fig. 4.10 User interface for Labview program

To control the galvo, the user inputs the amplitude for the driving wave for the *x* and *y* scanning dimensions, and the type of wave to drive the galvo system. For the *x* dimensions, a triangle wave is typically used since this is the fastest scanning path. For the *y* dimension, the galvo mirror is tilted after half the cycle of the triangle wave. The

frequency of the triangle wave is automatically calculated based off the requested pixel dwell time and the maximum frequency of the galvo scanning mirrors. An example of the waveforms sent to the galvo is shown in Fig. 4.11. Therefore, the 9 beams are raster scanned over the sample using these controls.

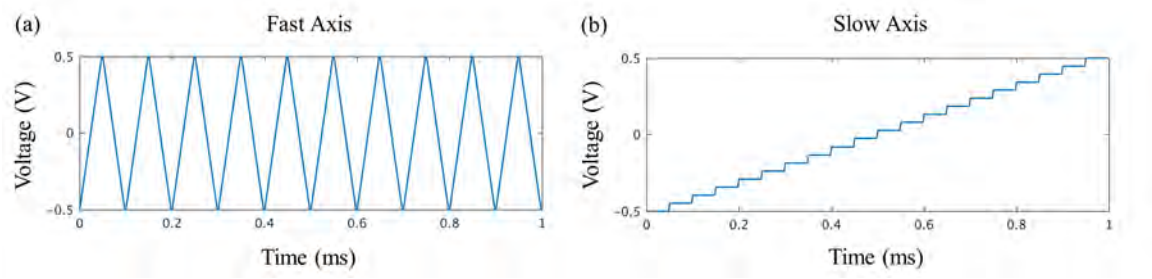

Fig. 4.11 (a) Triangle wave to drive fast axis of galvo. (b) Modified ramp to drive slow axis of galvo.

To control the translation stage, the user inputs the dimensions in the *x*, *y*, and *z* directions as well as the spacing in each of the direction. The system scans first in *z*, then translates in *y*, and then translates in *x*. The positions as read out from the motors and the progress in the scan for each dimension are displayed on the front panel of the user interface.

A 16-channel digitizer is used to convert the voltage readout from each channel of the PMT into a 14-bit digital number. The NI 5751 digitizer adapter module simultaneously samples 16 channels at 50 MS/s. The digitizer contains a fully programmable FPGA. Signal averaging over the pixel dwell time is done on the FPGAs to reduce the processing requirements by the computer. The FPGA averages the signal for the number of cycles that are within the pixel dwell time. One cycle for the FPGA is 25 ns. After averaging the signal from the PMT over the pixel dwell time, the digital value of the averaged voltage from the PMT is sent to the computer. The digital value from each channel is stored in a $3N_r \times 3N_c$ matrix. The position of each pixel in the matrix is dependent on the channel of the PMT and the position in the galvo scanning. After the galvo scan is completed, the

$3N_r \times 3N_c$ matrix is displayed, where all 9 images are shown adjacent to each other in a 3×3 grid.

Before the desired volume is scanned, the user should take a dark current measurement. Each channel in the digitizer has its own dark current reading associated with it, so this dark current should be removed before image acquisition of the volume of the sample. The user should block the laser beam and complete a scan to record the dark reading of the system after clicking the measure dark current button on the user interface. After one scan, the reading in each channel is averaged. This averaged dark current reading is then subtracted for every pixel after averaging the PMT signal on the FPGA.

Images from all 9 channels are shown together after every completion of the galvo scan. Note the $3N_r \times 3N_c$ matrix is stored in random-access memory (RAM) in the computer memory. Since volumes are going to be imaged, the total number of images acquired is large, so it is impossible to store all images in RAM. Therefore, after completion of every galvo scan, the $3N_r \times 3N_c$ matrix is saved to hard disk. To increase the speed of saving to the hard disk, the $3N_r \times 3N_c$ matrix is dumped from RAM as a binary file, instead of a more traditional image format. The binary files are read and analyzed in Matlab after completion of the entire volume scan.

In summary, the user inputs the number of rows and columns for each sub-image, the parameters for the galvo scan, and the positions to move the translation stage. After starting the scan, all 9 beams are scanned in a raster pattern by the galvo system. The digitizer with the FPGA averages the output voltage from each channel of the PMT over the pixel dwell time, and then sends the averaged output voltage to the computer. This averaged output voltage is placed in the correct position in a $3N_r \times 3N_c$ matrix. After the

galvo scan, the matrix is saved to hard disk. The translation stage then moves to the next position in the volume scan. This process is repeated for the total number of *x*, *y*, and *z* positions input by the user to scan the entire desired volume. After completion of the scan, the binary files at each position are then analyzed in Matlab to stitch and blend all 9 images together to create one large FOV image at each position.

4.2.7 Image Reconstruction Software

Matlab is used to take the binary files from each *z*-section of the volume and stitch the 9 sub-images together and blend the images from each sub-image, to create one blended large FOV image. First the background is estimated from each channel for every image in the volume and the background is subtracted for the image in that channel. The 9 images from each channel are then manually stitched together. A good object to use for calibration is a volume of fluorescent beads suspended in a transparent medium. After scanning a large volume, the centroids of the beads in the overlapping regions for each channel can be used as a guide to stitch the images together.

4.2.7.1 Image Blending

After the positions of the images for each channel are determined relative to each other, each channel needs to be blended with the adjacent channels. Figure 4.12(a) shows a simulation for the stitched image when each channel is not blended with the adjacent ones. Very sharp transitions between images from separate channels are visible in the stitched image. The test image in each channel has the intensity falling off radially. Figure 4.12(b) shows the image if the intensity of each channel is alpha blended with the adjacent channels in the overlapped region.

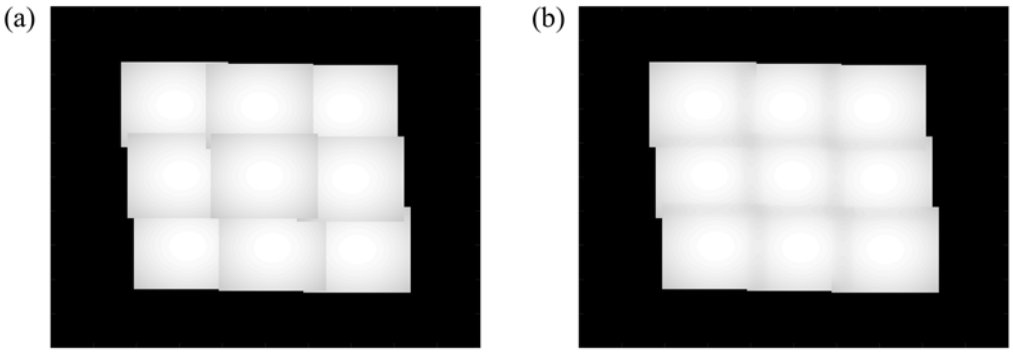

Fig. 4.12 (a) Stitched image from 9 channels with no blending. (b) Stitched image from 9 channels with alpha blending.

Suppose an image $I(x,y)$ needs to be blended with a composite image $C(x,y)$. The alpha blended image $B(x,y)$ for the overlapped region is

$$B(x,y) = \big(1-\alpha(x,y)\big)I(x,y) + \alpha(x,y)C(x,y) \tag{4.3}$$

where $\alpha(x,y)$ is the blending function with range [0,1]. The blending function is non-zero only in the regions where the image and composite image overlap. The process to blend all 9 images is as follows [78]:

1. Place the central image in center of an enlarged grid.

2. Place next image at appropriate position and find the overlapped region between image and the composite image.

3. Find the boundary of the overlapped region, and remove the borders of the overlapped region that encompass the new image from the boundary.

4. Compute the chessboard distance to the boundary in the overlapped region. The chessboard distance is $\max(|x_1-x_2|,|y_1-y_2|)$. Normalize the chessboard distance by the maximum distance. This is $\alpha(x,y)$.

5. Using Eq. (4.3) calculate the blended image.

6. Place blended image in the overlapped region of the composite image.

7. Repeat steps 2 – 6 for the all images.

Figure 4.13(a) shows an image of the overlapped region when the top left image is added to the composite blended image of the center image and the 4 images immediately adjacent to the central image. The boundary of the overlapped region is indicated as a red line in Fig. 4.13(a). The bottom and right borders in the boundary shown in Fig. 4.13(a) are removed in step 3 of the blending process, since these borders are the borders of the new image added to the composite image. Fig. 4.13(b) shows the results after blending the top left channel with the composite image. For this blended image the overlapped region is multiplied by $1-\alpha(x,y)$ and $\alpha(x,y)$, which are shown in Fig. 4.13(c) and 4.13(d), respectively.

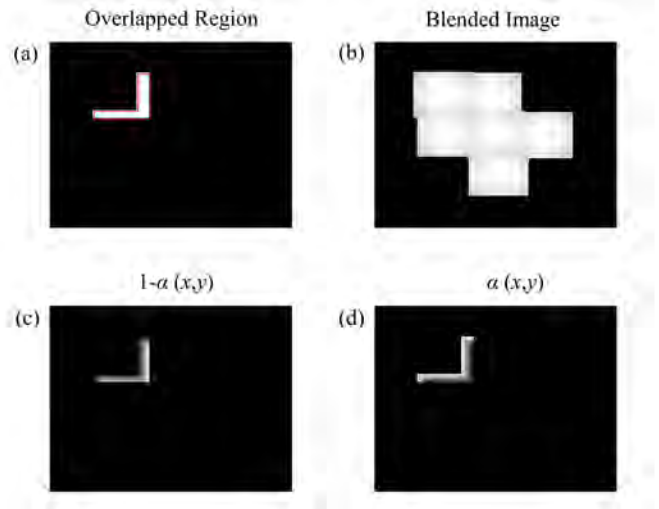

Fig. 4.13 (a) Overlapped region with boundary shown in red. (b) Blended image. (c) Image of $1-\alpha(x,y)$. (d) Image of $\alpha(x,y)$.

This blending method works well if the intensities in all channels respond linearly with respect to each other and there is no multiplicative difference between the intensities in each channel. Since the overlapped region is an image of the same object, the intensities in the overlapped region should be identical. One way to ensure the intensities are identical between the image added to composite and the composite image, is to minimize the intensity difference in the overlapped region, which results in

$$\min_{\beta,p}\left(\left|\left(\beta I(x,y)\right)^{p}-C(x,y)\right|^{2}\right) \tag{4.4}$$

where β and p are constants. To determine the best β and p, the intensity of pixels in the overlapped region between the image and composite image should have a large dynamic range. The blended image for the overlapped region is then calculated by

$$B(x,y)=\left(\beta\left(1-\alpha(x,y)\right)I(x,y)\right)^{p}+\alpha(x,y)C(x,y)\,. \tag{4.5}$$

Figure 4.14(a) shows a simulation if all 9 images are responding non-linearly and have a different multiplicative factor between them. Figure 4.14(b) shows the blended image using Eq. (4.5).

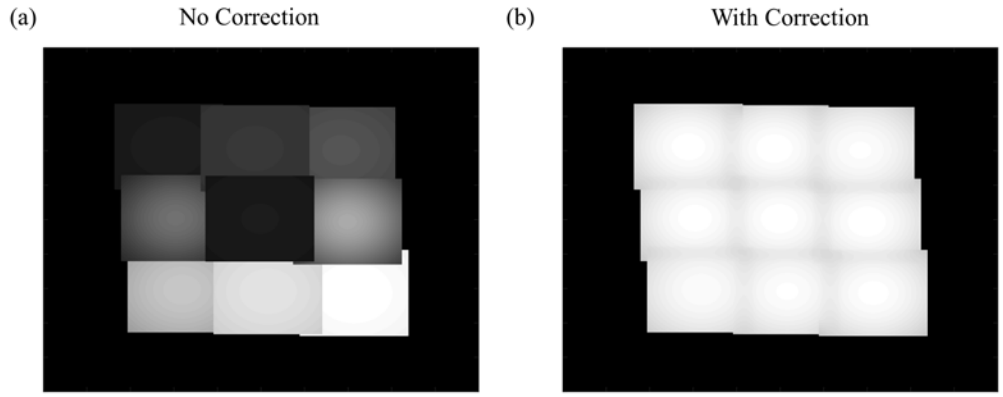

Fig. 4.14 (a) Stitched image when all channels respond non-linearly and with different multiplicative differences between all channels. (b) Stitched image after blending with non-linear correction.

To demonstrate the blending on images from the array confocal microscope, Fig. 4.15 shows the improvement using the blending process on an image from an axial slice of a rat brain. Figure 4.15(a) shows the images stitched together after background subtraction, but without blending or intensity normalization between channels. Figure 4.15(b) shows the blended image after blending and intensity normalization. The image is dark since the intensities of the pixels are normalized to the maximum intensity pixel value of the acquired value. Figures 4.15(c) brings out the lower intensity pixel values by applying a

gamma factor to the image in Fig. 4.15(b). Note that the sharp transitions and varying offset values in each channel are significantly reduced in the blended image.

Due to the large number of images acquired in imaging a volume and the large size of the images, the process of reconstructing all images in a volume can be long. Therefore, the code to stitch, blend and normalize the intensity is parallelized, so that multiple images are reconstructed in parallel by the computer. This parallelization greatly increases the speed required to reconstruct all images in a volume.

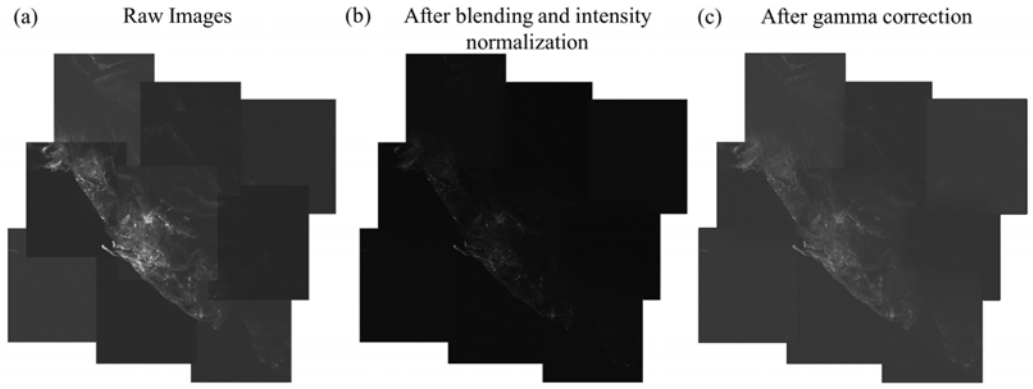

Fig. 4.15 (a) Raw image of brain after background subtraction. (b) Image of brain after blending and intensity normalization. (c) Image of blended and normalized brain with gamma factor to highlight lower intensity pixel values.

4.3 Experiments

The array confocal microscope is now used to image fluorescent microspheres and a cleared rat brain to demonstrate its imaging capabilities. Fluorescent microspheres are used to test the working distance and the resolution of the array confocal microscope.

4.3.1 Working Distance

To test the working distance, 0.5 μm diameter fluorescent microspheres are suspended in polydimethylsiloxane (PDMS), since the index of refraction is 1.4, which is close to the designed index of 1.45 for the objective. The cylindrical volume of fluorescent microspheres is placed in a container full of oil immersion liquid with an index of 1.45.

The volume is scanned 10 mm in depth with 3 μm spacing between consecutive images. Figure 4.16 shows the top view of the volume of microspheres. Microspheres are visible and distributed throughout all 9 channels. In some regions in the volume the microspheres did not disperse completely and the microspheres clumped together, which explains the larger objects in the volume. Figure 4.17(a) shows the *xz*-view and 4.17(b)-4.17(d) shows magnified *xz*-views in the top, middle, and bottom of the volume, respectively. The fluorescent microspheres are visible throughout the entire 10 mm, demonstrating the working distance of the objective is at least 10 mm. Figure 4.18(a) shows a tilted view and Fig. 4.18(b) shows the *yz*-view of the volume of fluorescent microspheres.

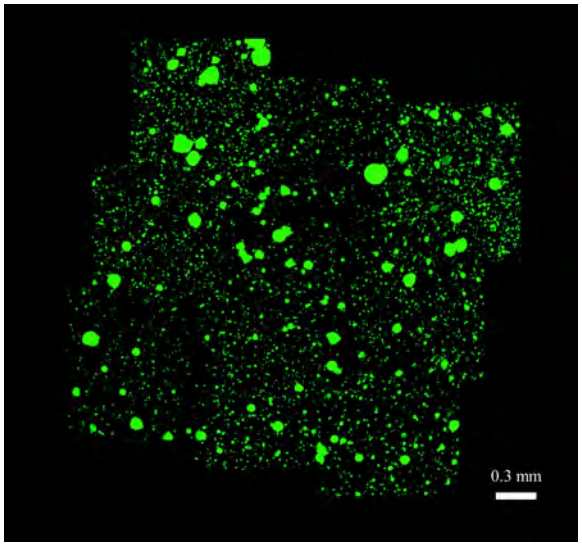

Fig. 4.16 Top view of volume of 0.5 μm fluorescent microspheres.

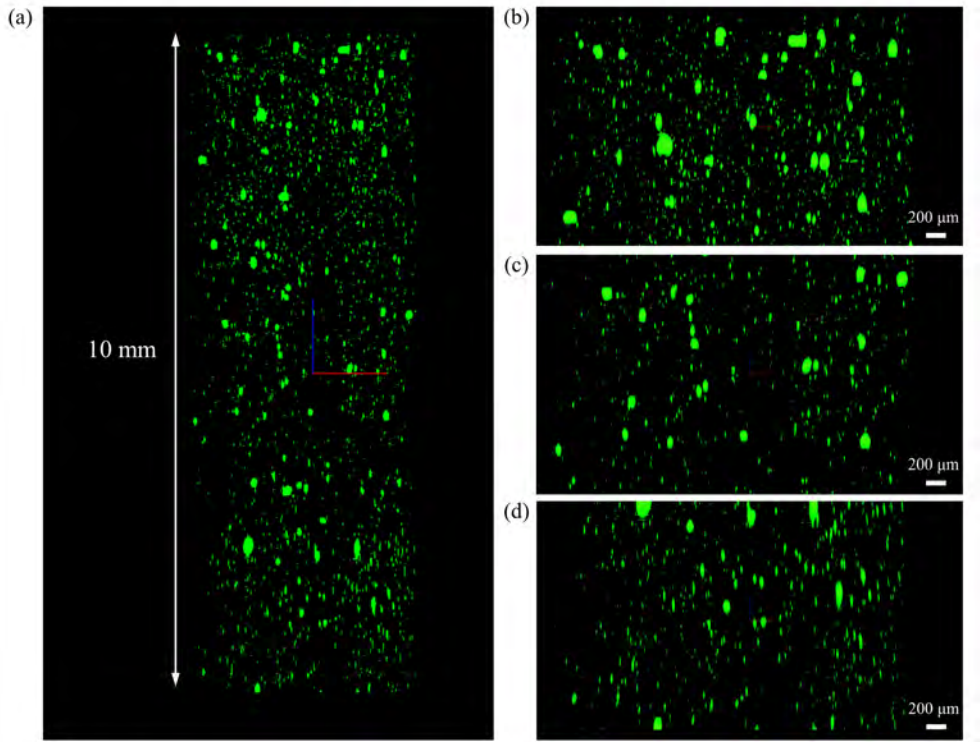

Fig. 4.17 (a) *xz*-view of volume of fluorescent microspheres. Magnified views of (b) top, (c) middle, and (d) bottom regions of the *xz*-view.

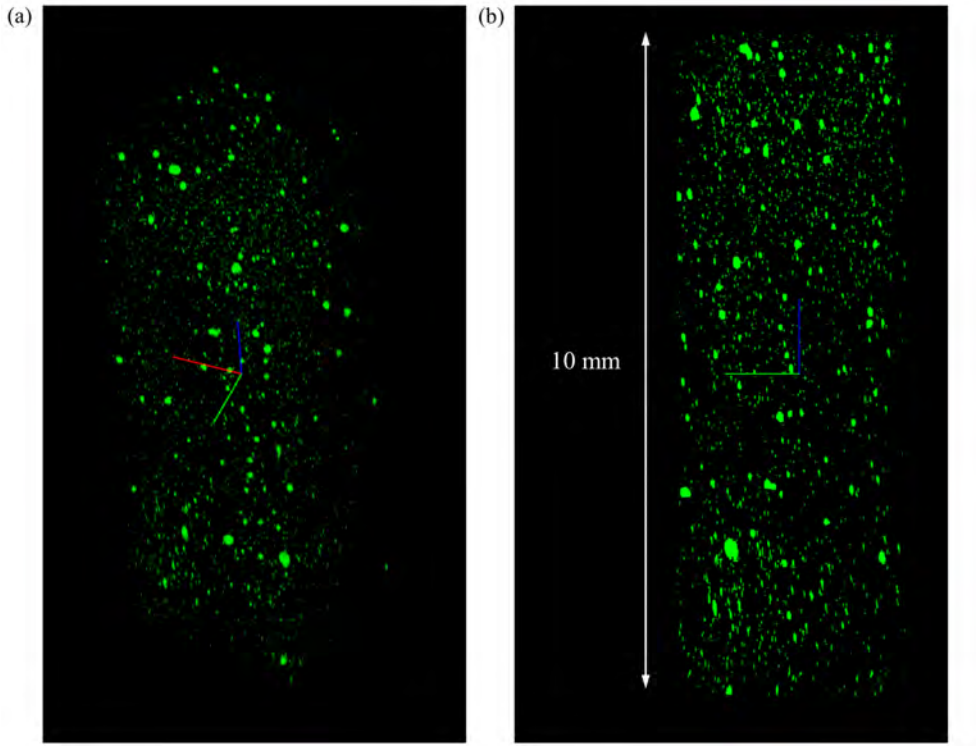

Fig. 4.18 (a) Tilted view of volume of fluorescent microspheres (b) *yz*-view of volume of fluorescent microspheres

4.3.2 Resolution

The resolution of the system is estimated by measuring the PSF in each of the 9 channels. The PSF is estimated by imaging 200 nm fluorescent microspheres [18]. The fluorescent microspheres are placed on a microscope slide and covered with a cover glass. The microscope slide is then placed in a container of oil immersion liquid. A cropped median-filtered image of one microsphere is shown in Fig. 4.19(a). A Gaussian function is fit to the image of the microsphere, which is used to calculate the FWHM of the PSF, as shown in Fig. 4.19(b). A cross-section of the bead is shown in the horizontal and vertical directions are shown in Fig. 4.19(c). The average FWHM for all 9 channels is 555 nm ± 88 nm. This corresponds to a resolution of 657 nm ± 104 nm. This resolution is slightly worse on average than the resolution of 628 nm for bright field microscopy and 511 nm for confocal microscopy. The confocal resolution is determined assuming the illumination and detection pinhole is a point. The detection pinhole in this system is 1.33 times the Airy disk radius, which is going to cause deviation from the confocal resolution. In practice, confocal microscopes give little lateral resolution improvement over bright field microscopy due to number of practical considerations [79]. Due to the complexity of this system, some misalignment between elements may also exist, which also reduces the resolution by introducing aberration into the system.

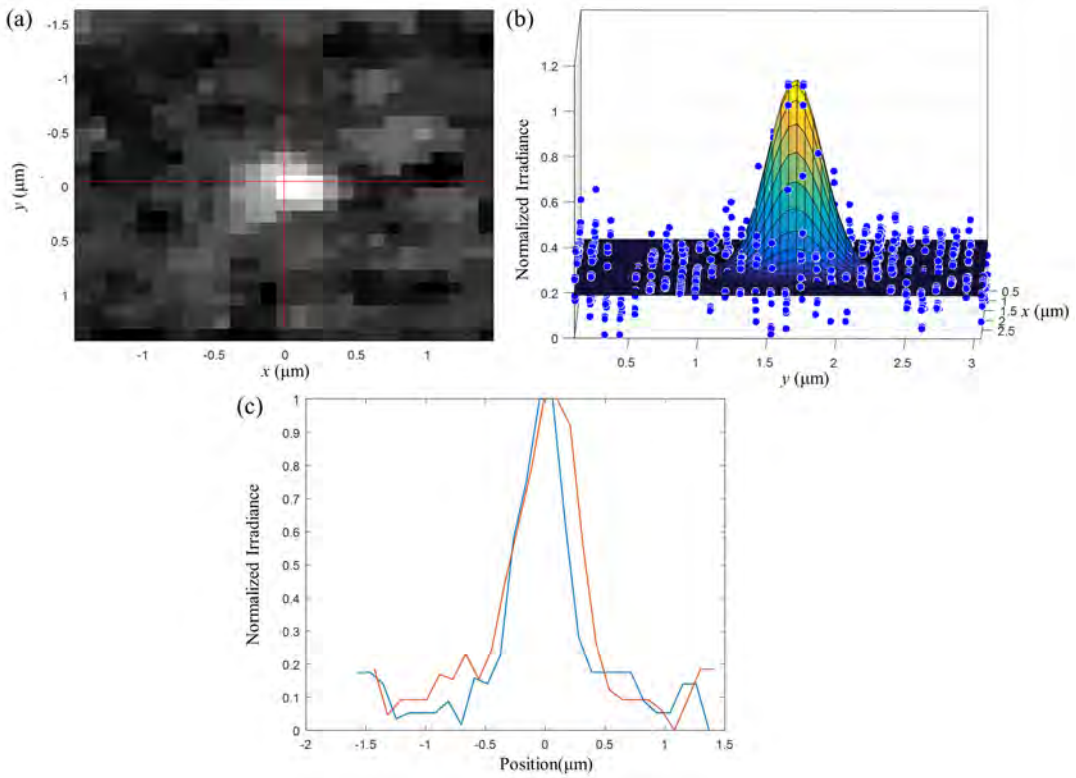

Fig. 4.19 (a) Image of 200 nm fluorescent microsphere, which approximates the PSF. (b) A Gaussian fit to the image of the microsphere. (c) A cross-section through lines displayed in (a).

4.3.3 Rat Brain

A rat brain is cleared using the brain clarification method CLARITY [74]. The cleared rat brain is immunostained with an antibody against an immediate early gene Arc. The cleared rat brain is imaged using the array confocal microscope. A photo of the cleared rat brain is shown in Fig. 4.20(a). Since the FOV of the array confocal microscope is too small to image the whole brain, one section of the brain is imaged in depth, and then the sample is translated to the next region. As shown in Fig. 4.20(a), the brain is not perfectly optically transparent. Due to scattering, the penetration depth of the confocal microscope is limited to approximately 500 μm for this sample. For each region, the cleared brain is imaged 2.55 mm in depth with an axial spacing of 50 μm. The brain is translated 8 times laterally, in a pattern similar to the one shown in Fig. 4.20(a). A top view of the 3D reconstruction of the fluorescence detected is shown in Fig. 4.20(b).

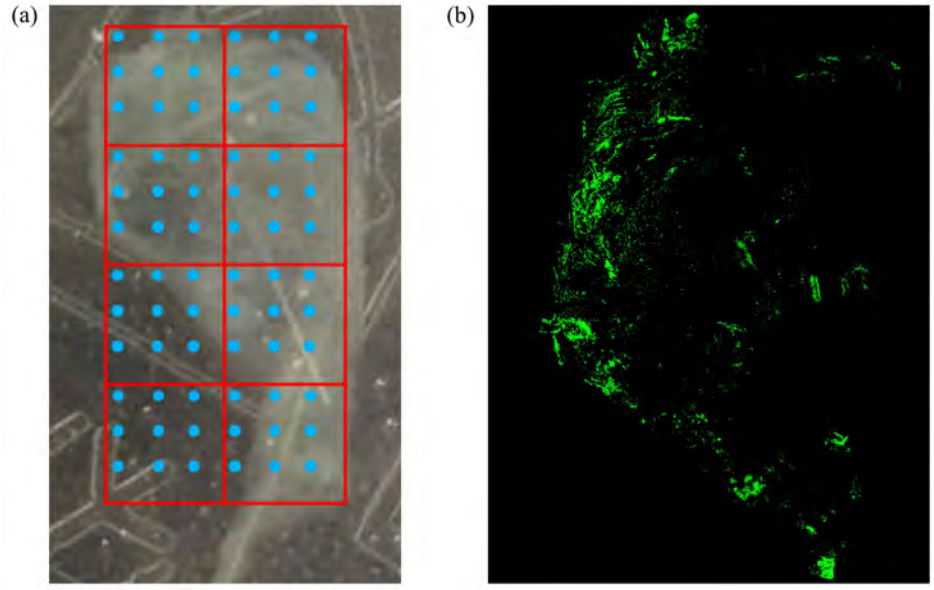

Fig. 4.20 (a) Photo of the cleared rat brain with a concept of the scanning pattern for the whole brain image. (b) A top view of the 3D reconstruction of the cleared rat brain over 2.2 mm in depth with a spacing of 50 μm.

Each channel scans approximately 1 mm × 1 mm in the horizontal and vertical directions. The images from the 9 channels are stitched and blended together as described in Sec. 4.2.7, to create an image with a 3 mm × 3 mm FOV. The pixel spacing is 0.62 μm. Four images of the cleared rat brain are shown in Figs. 4.21 and 4.22. All four images are from axial slices from different lateral regions in the cleared rat brain. Since the brain is not flat and there is a limited penetration depth with this sample, there is often not a detected signal in all 9 channels simultaneously.

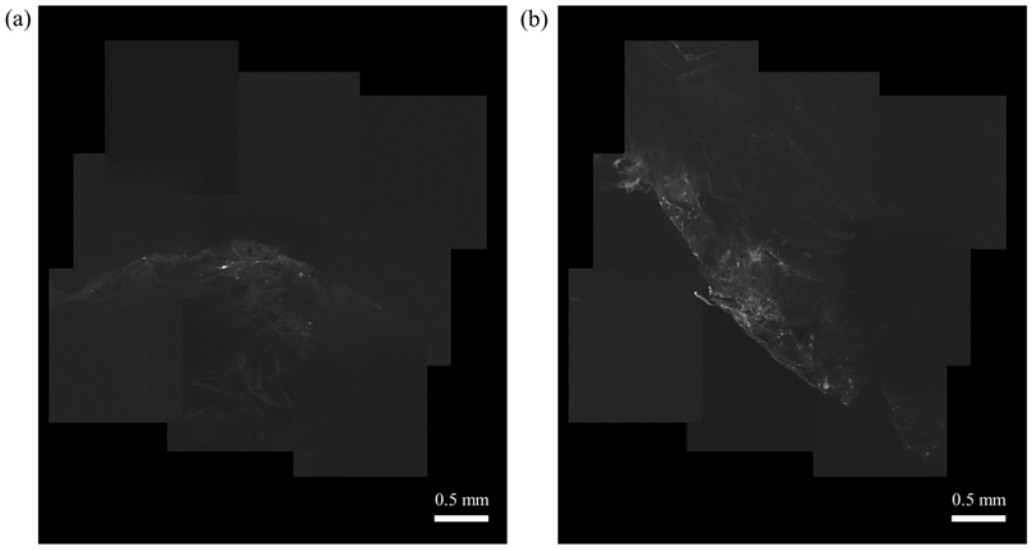

Fig. 4.21 (a,b) Image of an axial slice in the cleared rat brain.

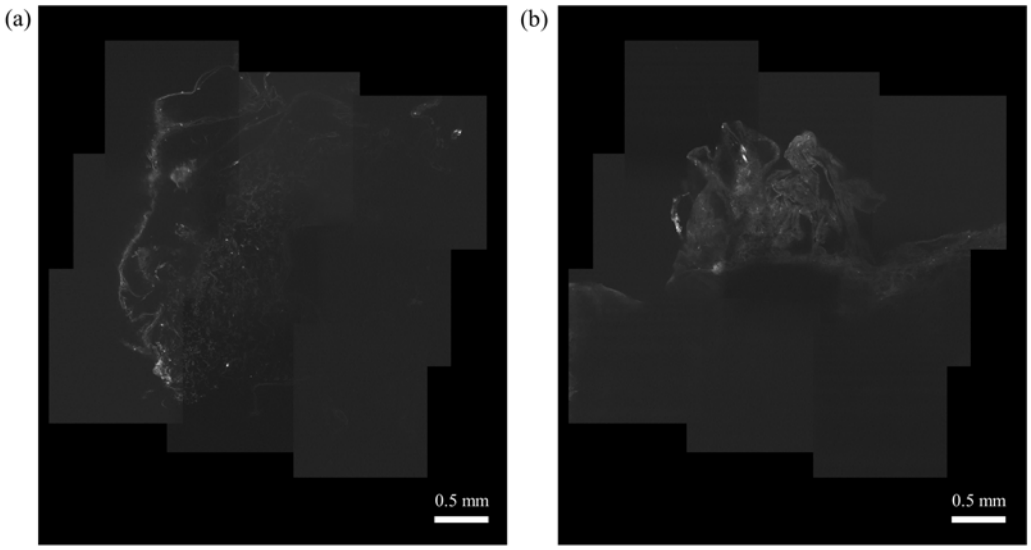

Fig. 4.22 (a,b) Image of an axial slice in the cleared rat brain.

The images from all 8 lateral regions are stitched together at one axial slice in Fig. 4.23(a) to demonstrate the potential for whole brain imaging. The dimensions of this image is approximately 7.0 mm × 10.8mm, with a pixel spacing of 0.62 μm throughout the image. Figure 4.23(b) and 4.23(c) shows magnified views of the regions in red indicated in Fig. 4.23(a). Since the array confocal microscope has subcellular resolution throughout the FOV, individual cells are visible in both Fig. 4.23(b) and Fig. 4.23(c).

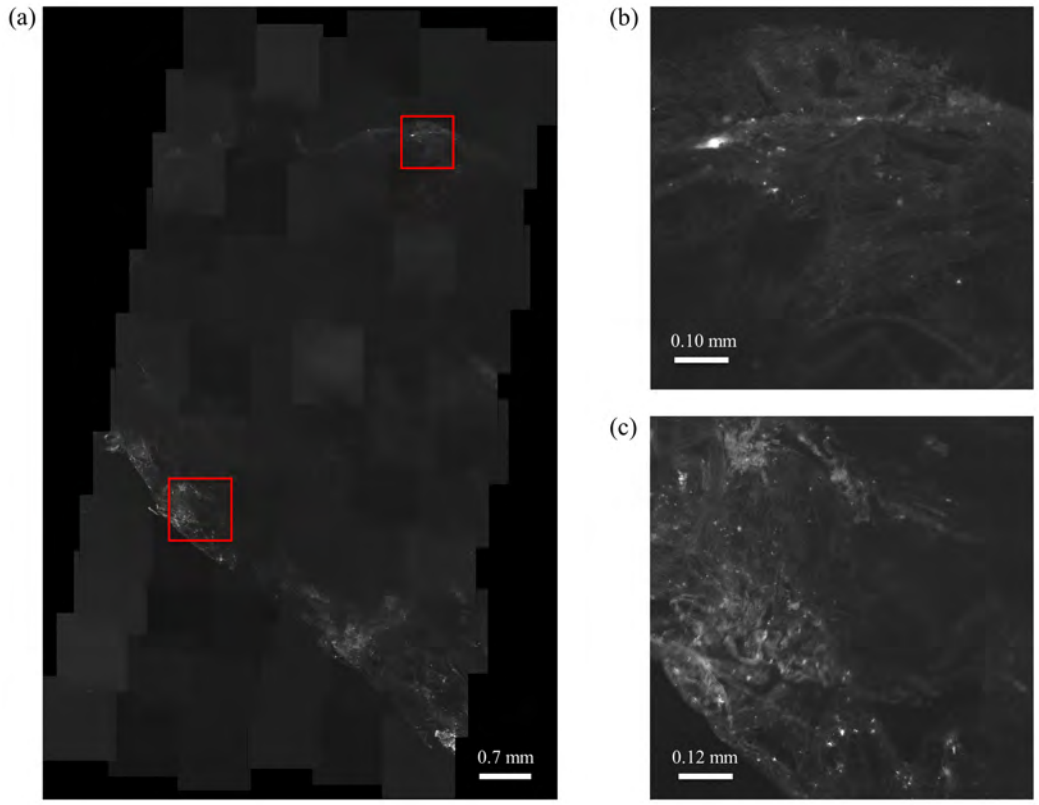

Fig. 4.23 (a) Images from all 8 lateral regions in the brain stitched together at one axial slice. (b) and (c) magnified views of the regions indicated in the red boxes in (a).

Now one region of the cleared rat brain is imaged with finer depth spacing. A depth of 2.2 mm is imaged in increments of 3 μm. Figure 4.24(a) shows the maximum projection image of this region in the cleared rat brain. Figure 4.24(b) shows a 3D reconstruction of this region in the brain, which shows the three-dimensional structure of the cells in this region of the brain. The hippocampus and the cortex are visible in this image.

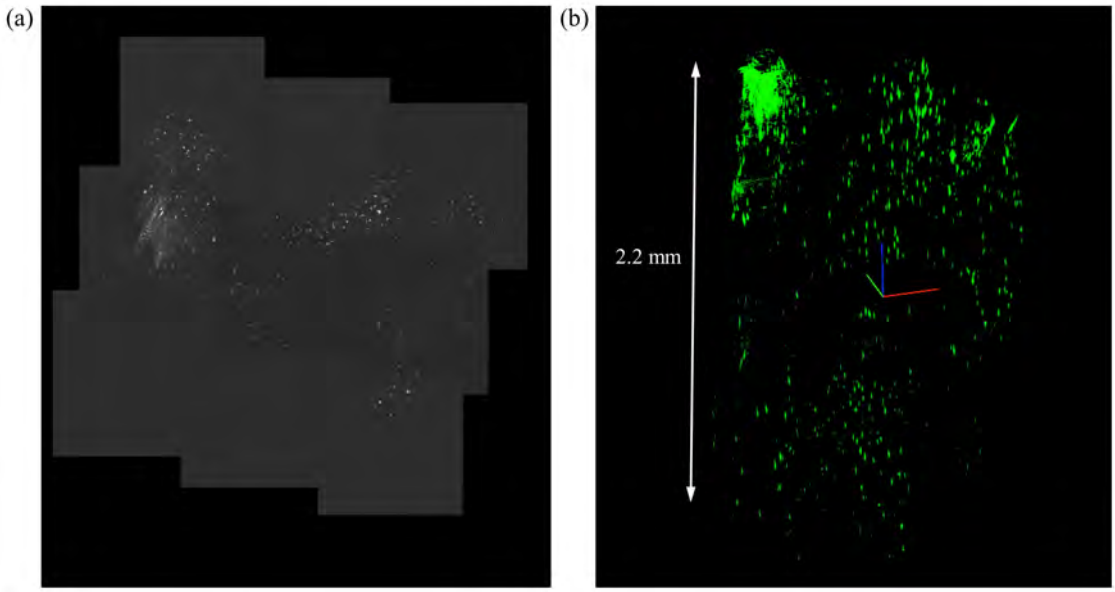

Fig. 4.24 (a) Maximum projection image of cleared rat brain over 2.2 mm in depth with 3 μm spacing in depth. (b) 3D reconstruction of the images from this region of the brain.

4.4 Conclusion

An array confocal microscope designed specifically for whole brain imaging of cleared brains is demonstrated. This microscope utilizes a custom 0.5 NA objective designed for a 3 mm × 3 mm FOV and 15.5 mm working distance. This array confocal microscope uses 9 beams to simultaneously scan the FOV to decrease the acquisition time by a factor of 9.

The array confocal microscope is built and tested. Fluorescent microspheres dispersed in a volume of PDMS showed the working distance is at least 10 mm in depth. A measurement of 200 nm fluorescent microspheres is used to estimate the PSF of the microscope in all 9 channels. The FWHM of the estimated PSFs are 555 nm ± 88 nm. Finally, a cleared rat brain is imaged to demonstrate the potential for whole brain imaging. At every axial position, the images from all 9 channels are stitched together to create an image with a 3 mm × 3 mm. In one experiment, the brain is translated 8 times laterally to image a 7 mm × 10.8 mm region of the brain with a lateral pixel spacing of

0.62 μm. The brain is imaged in axial steps of 50 μm to image a total depth of 2.55 mm. In another experiment, one lateral region of the brain is axially scanned 2.2 mm with an axial spacing of 3 μm. This experiment demonstrates the capability of the microscope to image a large volume at subcellular resolution, and reconstruct the three-dimensional layout of the cells in that region of the brain.

5. ANALYSIS OF GRATING DOUBLETS FOR ACHROMATIC BEAM-SPLITTING

5.1 Introduction

As shown in Chapters 2 and 4, beam-splitting gratings are instrumental components in an array confocal microscope, since they effectively and efficiently split an incident beam into N equal energy outgoing beams. However, the beam-splitting gratings designed in Chapters 3 and 4 are only designed for a single wavelength. The performance off the designed wavelength deteriorates significantly, which means array confocal microscopes using a single beam-splitting phase gratings are limited to a single excitation wavelength with good performance. The next two chapters analyze how grating doublets can be used to achromatize beam-splitting gratings for potential use in multi-wavelength array confocal microscopes. The results presented in this chapter are from one of my publications in Optics Express [80]. This publication proves the best height profiles for achromatic grating doublets have a constant first spatial derivative, and this publication shows a compensation method if one of the grating layers is fabricated with an incorrect height. This chapter shows some of these results. This chapter expands on this publication by demonstrating a method to reduce the height of the grating layers, since the height of the grating layers tend to be very large for achromatic grating doublets.

To demonstrate the performance change due to wavelength, the performance of a 1×5 continuous phase beam-splitting grating is simulated from 400 nm – 700 nm. The form of the phase function used to design beam-splitting continuous phase gratings into odd numbers of beams is described analytically by [63]:

$$\phi(x) = \tan^{-1}\left(\frac{Q(x,\boldsymbol{\alpha},\boldsymbol{\mu})}{P(x,\boldsymbol{\alpha},\boldsymbol{\mu})}\right), \tag{5.1}$$

$$P(x,\boldsymbol{\alpha},\boldsymbol{\mu}) = 1 + 2\sum_{j=1}^{M} \mu_j \cos(\alpha_j)\cos(jx), \tag{5.2}$$

$$Q(x,\boldsymbol{\alpha},\boldsymbol{\mu}) = 2\sum_{j=1}^{M} \mu_j \sin\left(\alpha_j\right)\cos\left(jx\right). \tag{5.3}$$

Note that the phase function ϕ is parameterized by the phase $\boldsymbol{\alpha}$ and amplitude $\boldsymbol{\mu}$ of corresponding sinusoids. The solution for a 1×5 beam-splitting grating with optimal efficiency into the first 5 modes has the following parameters: $\boldsymbol{\alpha}$ = (−π/2,π), $\boldsymbol{\mu}$ = (0.459, 0.899). The grating profile in radians is shown in Fig. 5.1(a). The grating is assumed to be fabricated using BK7 glass and is optimized for a wavelength of 486 nm. The standard deviation σ of the 5 desired modes as a function of wavelength is shown in Fig. 5.1(b). At 486 nm, all 5 desired output modes have equal energy, so σ is zero. As the wavelength varies from the optimized wavelength, the standard deviation of the desired 5 modes increases dramatically. At 700 nm, σ_{max}=0.165. The efficiency E of the 5 desired modes is shown in Fig. 5.1(c). The efficiency of each peak $|a_p|^2$ is shown in Fig. 5.1(d). The efficiency of the +1 mode and −1 mode change by the same amount, and likewise for the ±2 modes. At 700 nm, the maximum difference between modes is $\Delta|a_p|^2$ = 0.403. While this grating yields equal energy for the desired 5 modes at λ=486 nm, other wavelengths have a large deviation from the equal-energy condition. Due to the significant performance change with wavelength, this grating cannot be used with multiple wavelengths. Thus, a method to achromatize the grating must be used with multiple wavelengths.

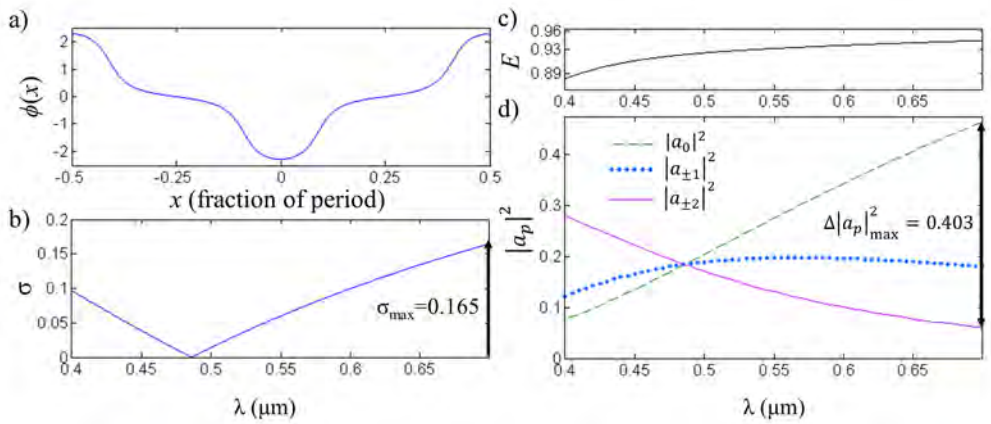

Fig. 5.1 (a) Transmitted phase in radians of achromatic 1×5 continuous phase beam-splitting grating, (b) Standard deviation σ and (c) total efficiency E of the 5 desired output modes as a function of wavelength. At 700 nm, σ_{max} = 0.165. (d) The efficiency of each output mode $|a_p|^2$ as a function of wavelength. At 700 nm, $\Delta |a_p|^2_{max}$ = 0.403.

There are several ways to split one laser beam into multiple beams with equal energy. For example, binary phase beam-splitting gratings utilize a pattern consisting of two phase levels [55-57]. Since these gratings are limited to only two phases, efficiency is relatively low. Dammann and Görtler were the first to address the problem of beam-splitting phase gratings [57], their class of binary phase grating solutions are now referred to as Dammann gratings. Continuous phase gratings were developed to further improve efficiency [57-59]. Using calculus of variations, Romero and Dickey found analytical expressions for the optimal continuous phase functions of one- and two-dimensional gratings that maximize energy into N desired output beams [63,64]. While these gratings theoretically provide optimal efficiency into N outgoing beams, the presence of fabrication errors can significantly degrade their performance. However, consideration of fabrication errors in the design can significantly improve tolerances [53]. Besides array confocal microscopy, other applications that benefit from beam-splitting gratings are parallel processing in laser machining and material processing, sensor systems, interferometry, communication systems, and image processing and gathering system [62].

As shown in Fig. 5.1, one disadvantage of single layer beam-splitting gratings is that they are only optimized for a single wavelength. If the wavelength deviates from the designed wavelength, uniformity of the output beams degrades. This degradation may be unacceptable for certain applications. An achromatic beam-splitting grating has wavelength-independent diffraction efficiency. By using multiple layers, a diffractive optical element (DOE) can be achromatized for multiple wavelengths [81,82]. The sensitivity of multilayer DOEs has been studied by analyzing the change in the diffraction efficiency of multilayer DOEs with respect to some common fabrication errors [83,84]. However, only DOEs that are designed to diffract all energy into one diffraction order were analyzed. By optimizing the levels of a single relatively thin multi-level phase grating, a grating can be achromatized at multiple wavelengths [85]. However, these designs typically have several closely spaced large discontinuities in the height profiles, which may make fabrication difficult.

In this chapter, a generalized achromatic design method is used to design multilayer achromatic beam-splitting grating doublets with equal energy output modes at two wavelengths. By studying the sensitivity of these achromatic grating doublets, it is shown that grating profiles with constant spatial derivatives perform significantly better than continuous grating profiles with respect to fabrication errors. Section 5.2 details the theory of achromatic beam-splitting grating doublet design under two possible fabrication errors: errors in the lateral alignment between the two grating layers and errors in the fabricated height in the grating layers. Section 5.3 simulates the performance of grating doublets to lateral shifts between grating layers for an achromatic continuous phase grating doublet and an achromatic Dammann grating doublet. Section 5.4 simulates the

performance of grating doublets to fabrication errors in the heights of each layer for an achromatic Dammann grating doublet. As will be shown, the heights of the grating layers in achromatic grating doublets can grow to be quite large for gratings, therefore, Section 5.5 describes a method to reduce the height of achromatic grating doublets.

5.2 Theory

A lossless, one-dimensional grating is characterized by periodic height function $h(x)$. In the absence of fabrication errors and using a thin grating approximation, the transmitted phase in air of a laser beam at position x is changed by an amount

$$\phi(x,\lambda) = h(x)2\pi(n(\lambda)-1)/\lambda, \tag{5.4}$$

where $n(\lambda)$ is the grating material refractive index at the wavelength of light λ. Note that the phase of the transmitted light on-axis is inversely proportional to the wavelength of light λ. For a grating illuminated by a unit-amplitude plane wave, the output beams, ignoring Fresnel losses, are characterized by the grating's Fourier coefficients:

$$a_p(\lambda) = \tfrac{1}{T}\int_{-T/2}^{T/2} e^{i\phi(x,\lambda)} e^{-i2\pi px/T} dx \text{ for } p = \pm 1, \pm 2, \ldots, \tag{5.5}$$

where T is the period of the grating. The Fourier coefficients determine energies and phases of the output beams [68].

The efficiency of each beam at wavelength λ is defined as $|a_p(\lambda)|^2$. A $1\times N$ vector $\eta[\phi(x,\lambda)]$ is defined that contains efficiencies of each desired output beam at wavelength λ,

$$\boldsymbol{\eta}_p[\phi(x,\lambda)] = \left|a_p(\lambda)\right|^2. \tag{5.6}$$

The efficiency of each output beam depends on the amount of fabrication error and the phase function of the grating. Total efficiency E of the grating at wavelength λ is

$$E[\phi(x,\lambda)] = \sum_{p=\boldsymbol{m}(1)}^{\boldsymbol{m}(N)} \left| a_p(\lambda) \right|^2, \tag{5.7}$$

where $\boldsymbol{m}$ contains the N desired output modes. Note that $\boldsymbol{m}$ is not limited to modes that are uniformly spaced. The remaining $(1-E)$ is spread into orders outside the desired N output modes. There may be significant energy in orders outside the desired range, so E might be less than unity. For a 1×5 beam-splitting grating, the total efficiency is the sum of N=5 desired output modes.

In the design algorithm, Fourier coefficients of an error-free uniform beam-splitting grating are subject to the constraint

$$\left| a_p(\lambda_i) \right|^2 = E / N, \tag{5.8}$$

which ensures that energies in all desired output beams are equal at wavelength λ_i. If there are fabrication errors or the wavelength changes, output beam energies in the desired range of p are not equal. The standard deviation σ is used to quantify the departure from the equal-energy output beam condition,

$$\sigma(\lambda) = s\left(\boldsymbol{\eta}[\phi(x,\lambda)]\right), \tag{5.9}$$

where s is the function for standard deviation. Since σ is the standard deviation of the beam efficiencies, σ is unitless.

Analogous to how a refractive achromatic doublet lens is designed to have the same focus for two wavelengths, an achromatic beam-splitting grating doublet is designed to produce σ = 0 at two wavelengths. In order for a beam-splitting grating to have equal energy output modes at two different wavelengths, two grating layers with different refractive indices are placed in series, as shown in Fig. 5.2 (a). To design achromatic grating doublets, start with the desired phase profile for a beam-splitting grating that is represented by $\phi(x)$ in radians. Achromatization with two grating layers sets $\phi(x)$ to be the

same at two wavelengths. The optical path lengths in units of microns for two desired wavelengths is given by:

$$OPL(x,\lambda_1) = (n_1(\lambda_1)-1)h_1(x) + (n_2(\lambda_1)-1)h_2(x) = \frac{\phi(x)}{2\pi}\lambda_1 \tag{5.10}$$

$$OPL(x,\lambda_2) = (n_1(\lambda_2)-1)h_1(x) + (n_2(\lambda_2)-1)h_2(x) = \frac{\phi(x)}{2\pi}\lambda_2 \tag{5.11}$$

where n_1 and n_2 are refractive indices for two different grating materials [81,82]. Using Eq. (5.10) and Eq. (5.11), the two unknown height profiles $h_1(x)$ and $h_2(x)$ are calculated. Note that this method can be expanded to optimize a grating for more than two wavelengths by adding more grating layers. All achromatic grating examples in this chapter use BK7 and SF5 for the grating layers, so the indices of refraction for these two materials as a function of wavelength is plotted from 400 nm – 700 nm in Fig. 5.2(b) and 5.2(c). Unless otherwise state, grating layer 1 is made from BK7 glass, and grating layer 2 is made from SF5 glass. Furthermore, the achromatic grating doublet examples in this chapter are 1×5 beam-splitting phase grating doublets that are optimized for λ_1 = 486 nm and λ_2 = 656 nm. At these wavelengths, n_{BK7}(486 nm) = 1.5224, n_{BK7}(656 nm) = 1.5143, n_{SF5}(486 nm) = 1.6875, and n_{SF5}(656 nm) = 1.6667.

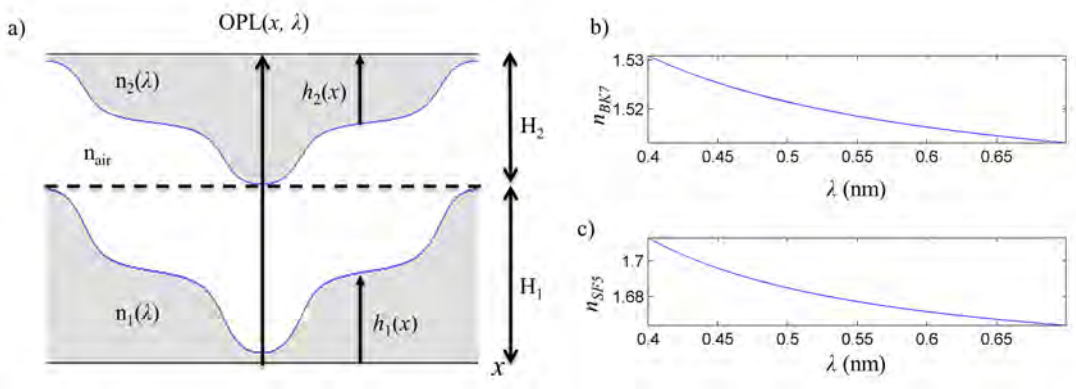

Fig. 5.2 (a) Diagram of the OPL for an achromatic grating. Index of refraction for (b) BK7 and (c) SF5 as a function of wavelength.

5.2.1 Lateral Shift between Grating Layers

Since the grating layers cannot be aligned perfectly, the effect of a lateral shift between two grating layers of an achromatic design is explored. First, assume the achromatic solution is known, where $\phi_1(x,\lambda)$ is the transmitted phase through grating layer 1 and $\phi_2(x,\lambda)$ is the transmitted phase through grating layer 2. When added together, $\phi_1(x,\lambda)$ and $\phi_2(x,\lambda)$ have a transmitted phase $\phi(x,\lambda)$ that produces equal energy output beams at two wavelengths λ_1 and λ_2. The electric field amplitude of the beam in the far field with uniform illumination is proportional to:

$$U_{equal} \propto \mathbf{F}\{e^{i(\phi_1(x,\lambda)+\phi_2(x,\lambda))}\}, \tag{5.12}$$

where **F** is the Fourier transform operator. A lateral shift Δx between grating layer 1 and grating layer 2 produces accumulated phase through both grating layers at a position x of

$$\phi_{total}(x,\lambda) = \phi_1(x,\lambda) + \phi_2(x+\Delta x,\lambda). \tag{5.13}$$

If Δx is small relative to the grating period, this expression is approximately

$$\phi_{total}(x,\lambda) \approx \phi_1(x,\lambda) + \phi_2(x,\lambda) + \phi_2'(x,\lambda)\Delta x, \tag{5.14}$$

where $\phi_2'(x,\lambda)$ is the first spatial derivative of $\phi_2(x, \lambda)$. The electric field amplitude in the far field of the two shifted grating layers with uniform illumination is proportional to:

$$U_{shift} \propto \mathbf{F}\left\{e^{i(\phi_1(x,\lambda)+\phi_2(x,\lambda)+\phi_2'(x,\lambda)\Delta x)}\right\} = \mathbf{F}\left\{e^{i(\phi_1(x,\lambda)+\phi_2(x,\lambda))}e^{i(\phi_2'(x,\lambda)\Delta x)}\right\}. \tag{5.15}$$

In order for the electric field to be unchanged by the relative shifts of the grating layers, either Δx=0 or $\phi_2'(x,\lambda)$=0. In either case, the multiplicative complex exponential is unity. By the convolution theorem, this expression is re-written as:

$$U_{shift} \propto \mathbf{F}\left\{e^{i(\phi_1(x,\lambda)+\phi_2(x,\lambda))}\right\} * \mathbf{F}\left\{e^{i(\phi_2'(x,\lambda)\Delta x)}\right\}, \tag{5.16}$$

where * is the convolution operator. The first Fourier transform yields U_{equal}, and the second Fourier transform is defined as U_{deriv}. In the far-field,

$$U_{shift} \propto U_{equal} * U_{deriv}. \tag{5.17}$$

If the shift is small, the electric field in the far-field is the convolution of U_{equal} with U_{deriv}. Since $\phi_2'(x,\lambda)$ is periodic, U_{deriv} is comprised of delta functions at integer multiples of the fundamental frequency. The solution that yields $U_{shift} = U_{equal}$ requires

$$e^{i\phi_2'(x,\lambda)\Delta x} = C, \tag{5.18}$$

where C is a constant.

To satisfy Eq. (5.18), $\phi_2'(x,\lambda)$ must be constant. The types of gratings that approximate this property are blazed gratings, Dammann gratings and multi-level phase gratings. Blazed gratings are ideally a sawtooth pattern, which has a constant non-zero first spatial derivative, except for the transitions at the end of every period. Dammann gratings and multi-level phase gratings have profiles with either binary or multi-level phases. Therefore, the grating profiles are a series of steps, where each step has a first spatial derivative equal to zero.

It makes intuitive sense that constant derivative grating profiles are the most resistant to lateral shifts between gratings. For example, as light is transmitted through two binary phase gratings shifted relative to each other, the only portion of the transmitted phase that deviates from the designed transmitted phase is near the transitions of the grating. All other portions of the transmitted phase have the designed phase. In comparison, if a continuous phase grating is shifted, the transmitted phase at every position x deviates from the designed transmitted phase after passing through both gratings. Therefore, the

far field pattern from a shifted continuous phase grating deviates significantly from its designed performance.

5.2.2 Error in Heights of Grating Layers

In addition to lateral shifts between the two grating layers, each grating layer may be fabricated with an error in its height. Start with the solution for an achromatic grating doublet, where $\phi_1(x, \lambda)$ is the transmitted phase through grating layer 1 and $\phi_2(x, \lambda)$ is the transmitted phase through grating layer 2. Together, $\phi_1(x, \lambda)$ and $\phi_2(x, \lambda)$ have a transmitted phase $\phi(x, \lambda)$ that creates equal energy output beams at two wavelengths λ_1 and λ_2. Assume there is a constant multiplicative fabrication error in the height, so the total phase transmission is $(1+\Delta)\phi(x, \lambda)$. The electric field amplitude in the far field in the presence of fabrication errors in the height is proportional to:

$$U \propto \mathbf{F}\left\{e^{i\left((1+\Delta_1)\phi_1(x,\lambda)+(1+\Delta_2)\phi_2(x,\lambda)\right)}\right\}, \tag{5.19}$$

where each grating layer has its own associated height error Δ_1 and Δ_2. To gain more insight, this expression is rewritten as

$$U \propto U_{equal} * \mathbf{F}\left\{e^{i\left(\Delta_1\phi_1(x,\lambda)+\Delta_2\phi_2(x,\lambda)\right)}\right\}. \tag{5.20}$$

For no change in U, the argument of the second Fourier transform, which is defined as U_Δ, must be constant,

$$U_\Delta = e^{i\left(\Delta_1\phi_1(x,\lambda)+\Delta_2\phi_2(x,\lambda)\right)} = C, \tag{5.21}$$

where C is a constant. Thus, the spatial dependence of the argument in Eq. (5.21) needs to be removed.

The transmitted phase is rewritten in terms of phase amplitude ϕ_{q0} and normalized base profile $b_q(x)$, where the subscript q indicates the layer of the doublet,

$$\phi_q(x,\lambda) = \phi_{q0}(\lambda) b_q(x), \tag{5.22}$$

where $b_q(x)$ ranges from [0,1]. Note achromatic grating doublet solutions have height profiles in which the two grating layers are related such that,

$$b_2(x) = 1 - b_1(x). \tag{5.23}$$

U_Δ written in terms of base function $b_1(x)$ is

$$U_\Delta = e^{i\left(\Delta_1\phi_{10}b_1(x,\lambda)+\Delta_2\phi_{20}\left(1-b_1(x,\lambda)\right)\right)} = e^{i\Delta_2\phi_{20}} e^{i\left(\Delta_1\phi_{10}b_1(x,\lambda)-\Delta_2\phi_{20}b_1(x,\lambda)\right)}, \tag{5.24}$$

The first term is a constant, so it can be factored out and ignored. To make the complex exponential unity, the spatially dependent terms are integer multiples of 2π for all x, so that

$$b_1(x)\left(\Delta_1\phi_{10}(\lambda) - \Delta_2\phi_{20}(\lambda)\right) = 2\pi m, \tag{5.25}$$

where m is an integer.

It is instructive to observe the conditions on Δ_1 and Δ_2 for which Eq. (5.25) is satisfied with different $b_1(x)$. For an arbitrary base profile function, the condition for which Eq. (5.25) is satisfied is given by

$$\Delta_2 = \frac{\phi_{10}(\lambda)}{\phi_{20}(\lambda)}\Delta_1 - \frac{2\pi m}{\phi_{20}(\lambda) b_1(x)}. \tag{5.26}$$

For m=0, this relationship holds for continuous and non-continuous grating profiles. If m is non-zero, Δ_2 is spatially dependent, since the offset is spatially dependent, which violates the assumption that Δ_2 is constant. However, suppose $b_1(x)$ is binary, so it only takes on values of 0 and 1. When $b_1(x)$ =0, Eq. (5.25) is satisfied, since 0=2πm when m=0. If $b_1(x)$ =1, Eq. (5.25) simplifies to

$$\Delta_1\phi_{10}(\lambda) - \Delta_2\phi_{20}(\lambda) = 2\pi m. \tag{5.27}$$

which is solved for Δ_2 as

$$\Delta_2 = \frac{\phi_{10}(\lambda)}{\phi_{20}(\lambda)}\Delta_1 - \frac{2\pi m}{\phi_{20}(\lambda)}. \tag{5.28}$$

Equation (5.28) is a linear relationship between Δ_1 and Δ_2 to produce no change to the designed σ in the far field pattern for binary grating doublets with height errors. Thus, there are lines of no change to the designed σ at integer multiples of 2π. The slope of this line for wavelength λ is

$$\beta(\lambda) = \frac{\phi_{10}(\lambda)}{\phi_{20}(\lambda)} = \frac{H_1\left(n_1(\lambda)-1\right)}{H_2\left(n_2(\lambda)-1\right)}, \tag{5.29}$$

where Eq. (5.4) is used for each grating layer. Note that both the slope and offset of this line are wavelength dependent. Even though the fabrication errors in height may combine to yield no change in the far-field pattern at one wavelength, the performance at another wavelength may change significantly.

5.3 Sensitivity to Lateral Shifts Between Grating Layers

This section simulates the performance for a 1×5 beam-splitting achromatic Dammann grating doublet and a 1×5 beam-splitting continuous phase achromatic grating doublet when the grating layers are misaligned by a lateral shift between layers.

5.3.1 Simulation of Performance of a Laterally Shifted Dammann Grating

Binary phase gratings have a first derivative of zero over nearly the entire period, which make them excellent candidates for achromatic beam-splitting grating doublets. The solution for an achromatic binary phase Dammann grating doublet that splits an incident beam into 5 equal energy output beams is shown in Fig. 5.3. The desired transmitted phase $\phi(x)$ has transitions between the high and low phases at x = (−0.471, −0.133, 0.133, 0.480)T, where T is the period of the grating, with a maximum phase of 2.993 radians

[62]. This grating is optimized to be achromatic at λ_1 = 486 nm and λ_2 = 656 nm. Grating layer 1 is made from BK7 glass and grating layer 2 is made from SF5 glass. Figure 5.3(a) shows the transmitted phase through the ideal combined grating layers, and Figs. 5.3(b) and 5.3(c) show height profiles of grating layers 1 and 2, respectively. The peak-to-valley heights are 11.29 μm and 8.24 μm for grating layers 1 and 2, respectively. Also, note that $h_2(x)$ is an inverted copy of $h_1(x)$, and is scaled by a different height. Figure 5.3(d) shows the efficiency E, which is nearly constant after λ = 486 nm. Figure 5.3(e) shows the efficiency of each of the 5 desired modes $|a_p|^2$ as a function of wavelength. After λ=460 nm, the efficiency of each mode remains relatively constant. A magnified portion of this wavelength range is shown in Fig. 5.3(f). This grating doublet has 5 nearly equal energy output beams from 486 nm – 700 nm, which makes it an excellent achromatic grating over this wavelength range with $\Delta|a_p|^2_{\max} = 0.010$ and $\sigma_{\max} = 0.004$ at λ=700 nm.

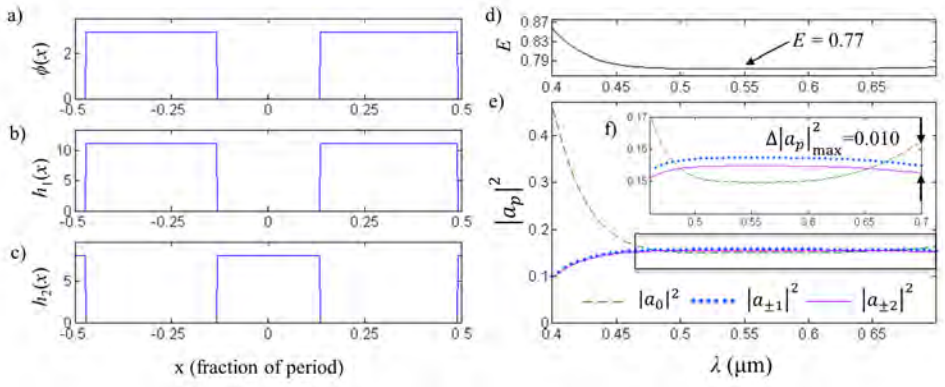

Fig. 5.3 Transmitted phase $\phi(x)$ in radians of an achromatic 1×5 beam-splitting Dammann grating doublet. (b) Height profile in microns for grating layer 1. (c) Height profile in microns for grating layer 2. Peak-to-valley heights are 11.29 μm and 8.24 μm, respectively. (d) Efficiency E of each of the desired 5 output modes from λ = 400 nm – 700 nm. (e) The efficiency of each output mode $|a_p|^2$ as a function of wavelength. (f) Magnified portion of $|a_p|^2$ from λ = 460 nm – 700 nm.

The transmitted phase $\phi_2(x, \lambda_1)$ is shown in Fig. 5.4(a) and $\phi_2'(x,\lambda_1)$ is shown in Fig. 5.4(b). $\phi_2'(x,\lambda_1)$ is zero, except at the transitions between the high and low phase of the profile. Figure 5.4(c) shows $|U_{deriv}|$ assuming $\Delta x = 0.01$, which is approximately a

delta function. Therefore, this grating is resistant to shifts between grating layers. The efficiency E and standard deviation σ are plotted as a function of the lateral shift in terms of the fraction of the grating period in Figs. 5.4(d) and 5.4(e), respectively. The difference in efficiency E between $\Delta x = 0$ and $\Delta x = 0.1$ is 0.05. Additionally, $\sigma_{max} \approx 0.04$ at $\Delta x = 0.1$ for $\lambda = 486$ nm. Figures 5.5(a) and 5.5(b) show efficiencies of each desired output mode as a function of the shift of the grating layers for λ_1 and λ_2, respectively. For small shifts there is very little change in the efficiencies of the desired modes, so the performance of this achromatic grating doublet is maintained, even in the presence of a lateral shift between the two grating layers.

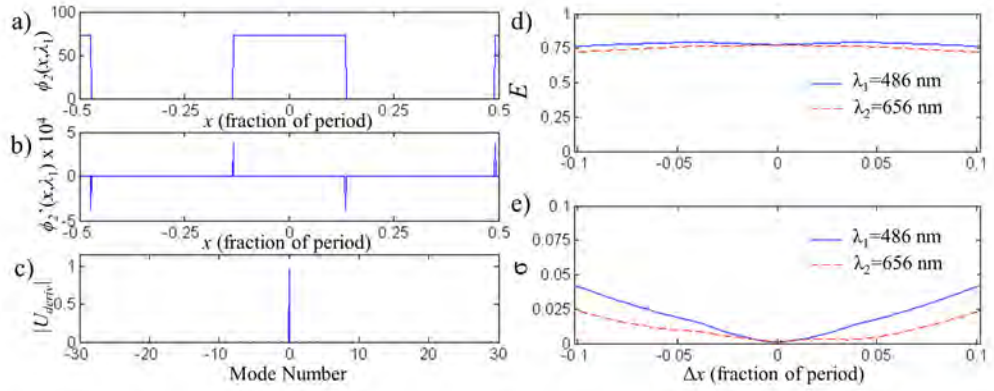

Fig. 5.4 (a) Transmitted phase in radians through $\phi_2(x,\lambda_1)$ of achromatic Damman phase grating doublet. (b) The spatial derivative $\phi_2'(x,\lambda_1)$ in units of radians/μm. (c) $|U_{deriv}|$. Note it is nearly a delta function. (f) Efficiency E and (e) standard deviation σ of desired 5 modes as a function of the lateral shift of grating 2.

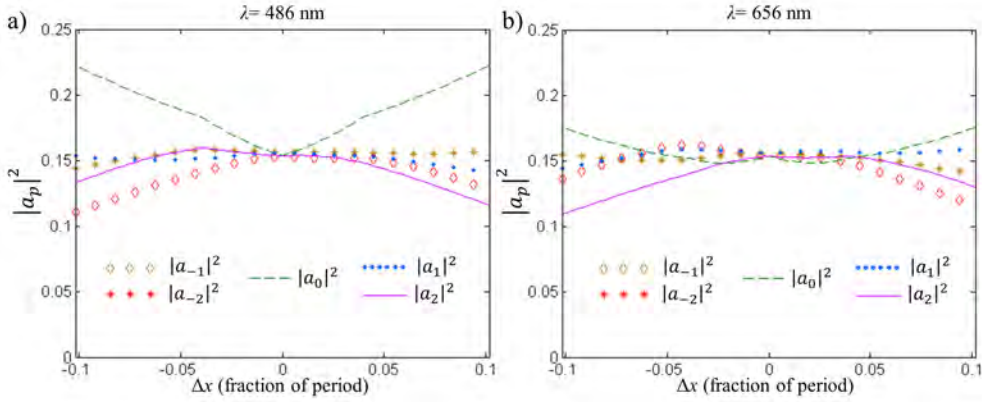

Fig. 5.5 (a) The efficiency of each output mode $|a_p|^2$ at (a) λ = 486 nm and (b) λ = 656 nm as a function of lateral shift of grating layer 2. At Δx = 0, $|a_0|^2 \approx |a_{\pm1}|^2 \approx |a_{\pm2}|^2$.

5.3.2 Simulation of Performance of a Laterally Shifted Continuous Phase Grating Doublet

A continuous phase grating doublet is analyzed to demonstrate problems associated with a lateral shift between two continuous grating layers. The solution for a 1×5 beam-splitting grating with optimal efficiency into the first 5 modes has the following parameters: $\boldsymbol{\alpha} = (-\pi/2, \pi)$, $\boldsymbol{\mu} = (0.459, 0.899)$. The designed phase profile for this grating is shown in Fig. 5.1(a). Grating layer 1 is made from BK7 glass, and grating layer 2 is made from SF5 glass. After achromatization, grating layer 1 has a peak-to-valley height of H_1 of 17.24 μm and grating layer 2 has a peak-to-valley height H_2 of 12.58 μm. Figure 5.6(a) shows the transmitted phase $\phi_2(x,\lambda)$ at λ_1. Figure 5.6(b) shows $\phi_2'(x,\lambda_1)$. Figure 5.6(c) shows $|U_{deriv}|$ assuming $\Delta x = 0.01$, where units of Δx are fractions of the period. Note that $|U_{deriv}|$ exhibits significant energy in almost all modes. When U_{deriv} is convolved with U_{equal}, a low-amplitude copy of U_{equal} is centered at each of the modes shown in Fig. 5.6(c). This solution yields extremely poor performance, even with the small shift of Δx=0.01. Figure 5.6(d) shows the total efficiency E and Fig. 5.6(e) shows the standard deviation σ of the 5 desired output beams plotted as a function of the relative shift between gratings at the two design wavelengths. For $|\Delta x| > 0.008$, efficiency E drops to below half its designed value, which indicates the majority of the energy in these modes are distributed to higher, undesired modes. The difference in E between $\Delta x = 0$ and $\Delta x = 0.1$ is 0.86. The maximum value of σ is $\sigma_{max} \approx 0.1$ at $\Delta x \approx 0.004$, which is a significant departure from the design value. Figures 5.7(a) and 5.7(b) show $|a_p|^2$ for the continuous phase grating for λ_1 and λ_2, respectively. After $|\Delta x| > 0.01$, the efficiency of each output mode is below half its designed value. The performance of this grating

deteriorates significantly with small lateral shifts. Due to the high sensitivity to lateral shifts between the grating layers, continuous phase gratings are ill-suited for achromatic grating doublets. Continuous phase achromatic grating doublets require nearly perfect alignment, which may be extremely difficult to achieve during fabrication.

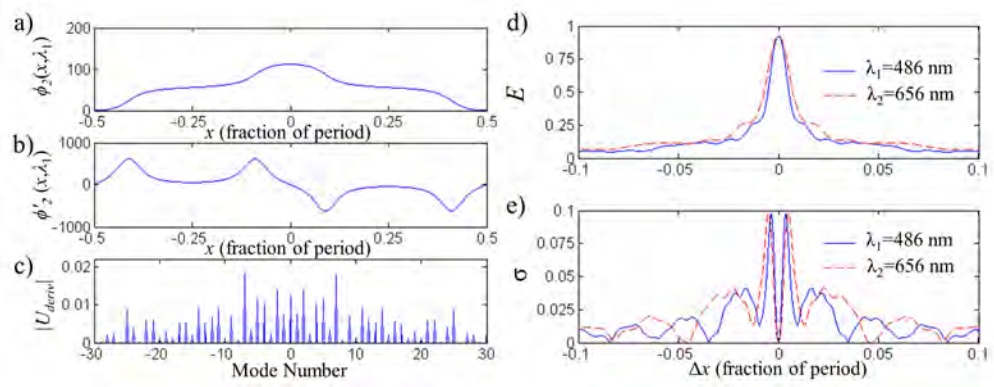

Fig. 5.6 (a) Transmitted phase $\phi_2(x,\lambda)$ in radians through the grating layer 2 of 1×5 achromatic continuous phase beam-splitting grating doublet. (b) The spatial derivative $\phi_2'(x,\lambda_1)$ in units of radians/μm. (c) $|U_{deriv}|$. Note nearly all modes are non-zero. (d) Efficiency E and (e) standard deviation σ of desired 5 modes as a function of the lateral shift of grating layer 2.

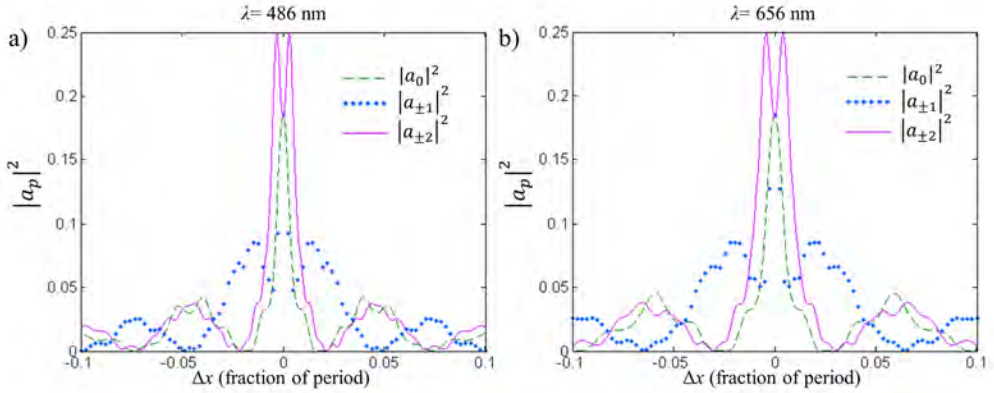

Fig. 5.7 Efficiency of each output mode $|a_p|^2$ at (a) λ = 486 nm and (b) λ = 656 nm as a function of lateral shift of grating layer 2. At Δx = 0, $|a_0|^2 \approx |a_{\pm1}|^2 \approx |a_{\pm2}|^2$.

5.3.3 Comparison between a Laterally Shifted Dammann Grating Doublet and Continuous Phase Grating Doublet

Greisukh *et al* demonstrated scalar diffraction theory reliably models two-layer phase-relief diffraction structures when the ratio of the structure period to the height of the diffraction structures is greater than 2.5 [86]. Based off this result, scalar diffraction

theory is reliably accurate for periods greater than 28 μm and 43 μm for the achromatic Dammann grating doublet and the continuous phase achromatic grating doublet, respectively. To maintain $\sigma<0.04$ for the achromatic continuous phase grating doublet at λ=486 nm with a 43 μm period, the grating layers have to be aligned with $\Delta x < 0.002$, which corresponds to a lateral shift less than 0.086 μm. In contrast, the achromatic Dammann grating doublet with a 43 μm period has to be aligned within 4.3 μm to maintain $\sigma<0.04$ at λ=486 nm. For this example, the tolerance for alignment is 50 times smaller for the achromatic continuous phase grating doublet, which makes alignment extremely difficult.

A summary of the results for the achromatic continuous phase grating doublet and the achromatic Dammann grating doublet is shown in Table 5.1. Note the achromatic Dammann grating doublet significantly outperforms the achromatic continuous phase grating doublet in all cases. Therefore, gratings that have a height profile where the spatial derivative is constant over approximately the entire period like Dammann gratings, multi-level gratings, and blazed gratings minimize sensitivity to relative shifts between the two grating layers. While continuous phase beam-splitting gratings are superior for single-wavelength designs, since they can yield higher efficiencies and better uniformity than Dammann and multi-level gratings, their sensitivity to lateral shifts between grating layers makes it extremely difficult to fabricate an achromatic continuous phase grating doublet with good performance. Dammann gratings, multi-level gratings, and blazed gratings are superior for fabrication of achromatic beam-splitting grating doublets.

Table 5.1. Summary of 1×5 achromatic grating doublet results

	λ=486 nm – 700 nm Δx = 0		λ_1=486 nm Δx = 0.004		λ_2=656 nm Δx = 0.004	
	σ_{max}	$\Delta \mid a_p \mid^2_{max}$	σ	$\Delta \mid a_p \mid^2$	σ	$\Delta \mid a_p \mid^2$
Continuous phase grating doublet	0.020	0.047	0.097	0.187	0.089	0.173
Dammann grating doublet	0.004	0.010	0.001	0.003	0.002	0.004

5.4 Simulation of Height Fabrication Errors for Achromatic Dammann Grating Doublets

The performance change due to height errors in both layers of the achromatic Dammann grating doublet described in Sec. 5.2.2 is simulated in this section. Figure 5.8(a) shows σ as a function of Δ_1 and Δ_2 for λ=486 nm. Note that σ is periodic, as predicted by Eq. (5.27). Furthermore, the minima of σ are located along the lines given by Eq. (5.28), shown as solid lines in Fig. 5.8(a). Assume grating layer 1 is fabricated with the incorrect height after fabrication and is measured to have multiplicative height error Δ_1. Instead of trying to fabricate grating layer 2 at the designed value $\Delta_2 = 0$, σ is minimized if the height of grating layer 2 is calculated using the line given by Eq. (5.28).

As predicted by Eqs. (5.27) and (5.28), σ is also minimized when *m* is non-zero. To demonstrate the periodic nature of σ, the line perpendicular to the lines given by Eq. (5.28) is shown as a dashed line in Fig. 5.8(a). For λ=486 nm, σ is periodic along the dashed line with a period T_{486} as shown in Fig. 5.8(b). However, this period is wavelength dependent. Figure 5.8(b) also shows σ perpendicular to the lines given by Eq. (5.28) for λ=656 nm, which has a larger period T_{656}. When *m* is zero in Eq. (5.27), the minima of σ for different wavelengths overlap. However, when *m* is non-zero, the minima of σ for different wavelengths do not overlap. Thus, when averaged over multiple wavelengths, the performance of the doublet is reduced when *m* is non-zero.

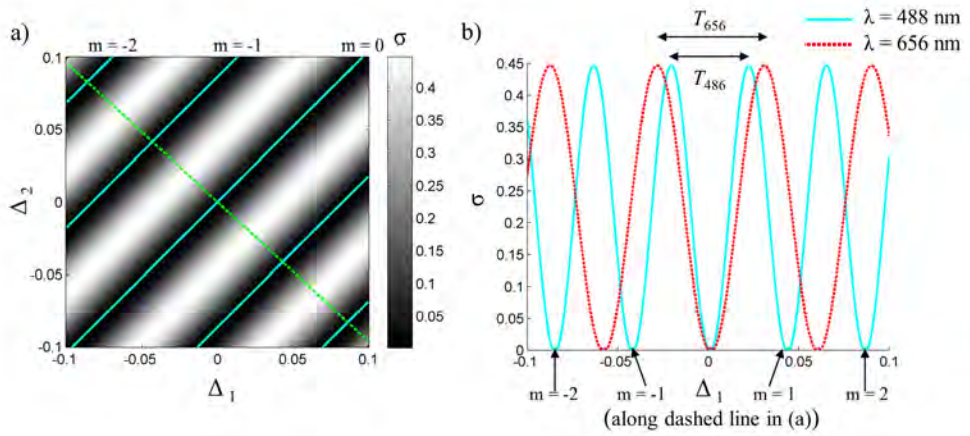

Fig. 5.8 (a) Standard deviation σ of 5 desired output modes of achromatic Dammann grating doublet as a function of height fabrication errors Δ_1 and Δ_2 at λ = 486 nm. The lines given by Eq. (5.28) for λ = 486 nm are plotted as solid lines. The line perpendicular to the solid lines is plotted as a dashed line. (b) σ along dashed line in (a) for λ = 486 nm (solid line) and λ = 656 nm (dashed line).

Since the achromatic Dammann grating doublet has excellent performance from λ=486 nm – 700 nm, Fig. 5.9(a) shows σ averaged over this wavelength range as a function of Δ_1 and Δ_2. The multiplicative height error combinations that minimize σ_{avg} are shown along the solid line in Fig. 5.9(a). Even with errors from Δ_1=[−0.1,0.1], σ_{avg} only changes by a maximum of 5×10^{-5} along this line. In comparison, suppose Δ_2 is fabricated and has a value along the dashed line shown in Fig. 5.9(a). Figure 5.9(b) shows σ_{avg} along the dashed line as a function of Δ_1. Along this line, σ_{avg} oscillates, but does not return to the minimum value at Δ_1=0 since Eq. (5.27) cannot be satisfied for multiple wavelengths when *m* is non-zero.

As an example, suppose grating layer 1 is fabricated with Δ_1 = 0.05. If grating layer 2 is fabricated with Δ_2=0, σ_{avg} = 0.429, which is indicated as the red square in Fig. 5.9(a). However, if grating layer 2 is calculated using Eq. (5.28) it is fabricated with Δ_2=0.049, which is indicated as the yellow circle in Fig. 5.9(a). At this Δ_2, σ_{avg} reduces to 0.002, which is a 215 times improvement. Therefore, the performance of the achromatic

Dammann grating doublet is improved significantly over a large wavelength range if the height of grating layer 2 is calculated taking Δ_1 into account.

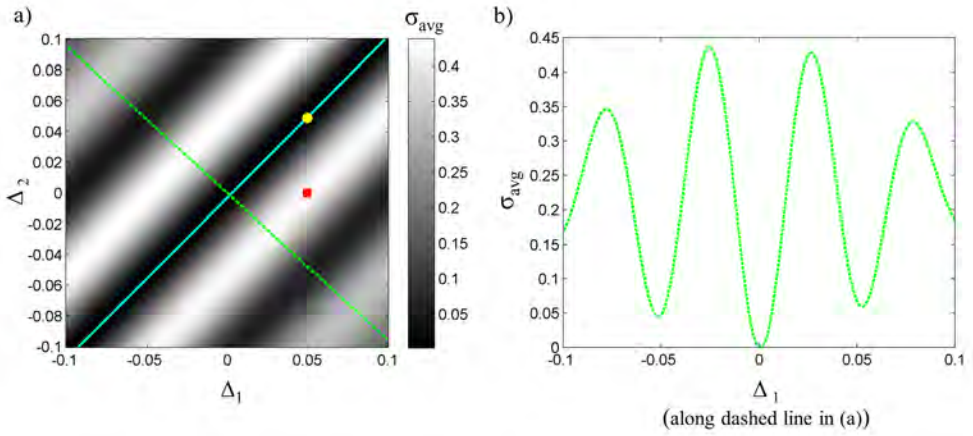

Fig. 5.9 (a) Standard deviation σ of 5 desired output modes of achromatic Dammann grating doublet as a function of height fabrication errors Δ_1 and Δ_2 averaged over λ = 486 nm – 700 nm. Solid line is height combinations that minimize σ_{avg}. Dashed line is the perpendicular to the solid line. At red square (0.05,0), σ_{avg} = 0.429. At yellow circle (0.05,0.049), σ_{avg} = 0.002. (b) Average σ from 486 nm – 700 nm along dashed line in (a).

5.5 Height-Reduced Achromatic Grating Doublets

One major disadvantage of achromatic grating doublets is the large heights of each grating layer. For the 1×5 achromatic Dammann grating, the heights of the two layers are 11.29 μm and 8.24 μm, respectively. These large heights make fabrication difficult. Suppose instead of designing a binary grating with a designed phase of ϕ_0, a phase of $\phi_0+2\pi m$ is used, where m is an integer. The transmitted electric field through the grating is

$$u_{grating}(x) = Ce^{i(\phi_0+2\pi m)b(x)} = Ce^{i\phi_0 b(x)}e^{i2\pi m b(x)} = Ce^{i\phi_0 b(x)}. \quad (5.30)$$

where C is a constant. Therefore, the addition of an integer multiple of 2π to the phase of a binary grating is an equivalent transmitted electric field. Instead of using Eqs. (5.10) and (5.11) to calculate the heights of the grating layers, the phase plus an integer multiple

of 2π is used to reduce the height of the grating layers. The OPL at two designed wavelengths is

$$OPL(x,\lambda_1) = (n_1(\lambda_1)-1)h_1(x) + (n_2(\lambda_1)-1)h_2(x) = \frac{b(x)\left(\phi_0 + 2\pi m\right)}{2\pi}\lambda_1 \tag{5.31}$$

$$OPL(x,\lambda_2) = (n_1(\lambda_2)-1)h_1(x) + (n_2(\lambda_2)-1)h_2(x) = \frac{b(x)\left(\phi_0 + 2\pi p\right)}{2\pi}\lambda_2 \tag{5.32}$$

where *m* and *p* are integers. This method can be used to search for a grating solution with grating layers with reduced heights.

The phase profile for the 1×5 achromatic Dammann grating used in the Sec. 5.3 is used to design an achromatic grating for λ = 486 nm and 656 nm with reduced depth. Once again, grating layer 1 is made from BK7 glass and grating layer 2 is made from SF5 glass. The heights of both grating layers are solved using Eq. (5.31) and Eq. (5.32) for the range of integers m = [−100,100] and p = [−100,100]. The logarithm of the sum of the two heights is shown in Fig. 5.10. By using different *m* and *p*, the sum of the heights changes dramatically. The minimum summed height is shown as a red dot in Fig. 5.10. The minimum summed height does not occur when m=0 and p=0, which has a summed height of 19.53 μm. The minimum summed height occurs when m=−4 and p=−3 with a summed height of 3.62 μm. In fact, there are 12 solutions with summed heights less than m=0 and p=0 over this range of integers for this grating example.

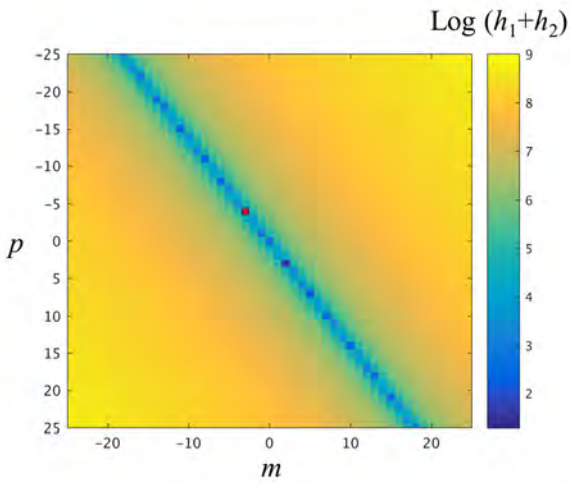

Fig. 5.10 Log(h_1+h_2). The position of the minimum summed height is shown as a red dot at m=−4 and p = −3.

Figure 5.11(a)-(d) shows the performance of the achromatic grating when m=0 and p=0. The heights of grating layer 1 and 2 are 11.29 μm and 8.24 μm, respectively, as shown in Fig. 5.11(a) and 5.11(b). The efficiency of this grating doesn't vary much from 486 nm through 700 nm as shown in Fig. 5.11(c). Figure 5.11(d) shows the efficiency in each of the first five diffraction orders is approximately equal from 486 nm through 700 nm. Figures 5.11(e) and 5.11(f) show the height profiles for the grating layers when m=−4 and p=−3. This is the solution with the minimum summed height. Note grating layer 1 is now only 0.64 μm, which can easily be fabricated using traditional lithographic methods, and grating layer 2 is 2.98 μm. Note that the efficiency is no longer constant over the wavelength range from 486 nm through 700 nm as shown in Fig. 5.11(g). Figure 5.12(h) shows the efficiency of the first five diffraction orders, where the efficiencies of ±1 and ±2 orders overlap in Fig. 5.11(h). The performance of this achromatic grating is very different than the m=0 and p=0 solution. Since there is a phase difference of 2π between the phase at λ = 486 nm and 656 nm, the transmitted phase through the grating passes through a full 2π phase from λ = 486 nm through 656 nm. The grating performs as designed performance at λ = 486 nm and 656 nm, however, the performance between

these two wavelengths is significantly degraded. While this method reduces the depth of these gratings, if $m \neq p$ there is a multiple of 2π phase difference between the phase at λ_1 and λ_2, the performance of the grating between these two wavelengths is significantly degraded.

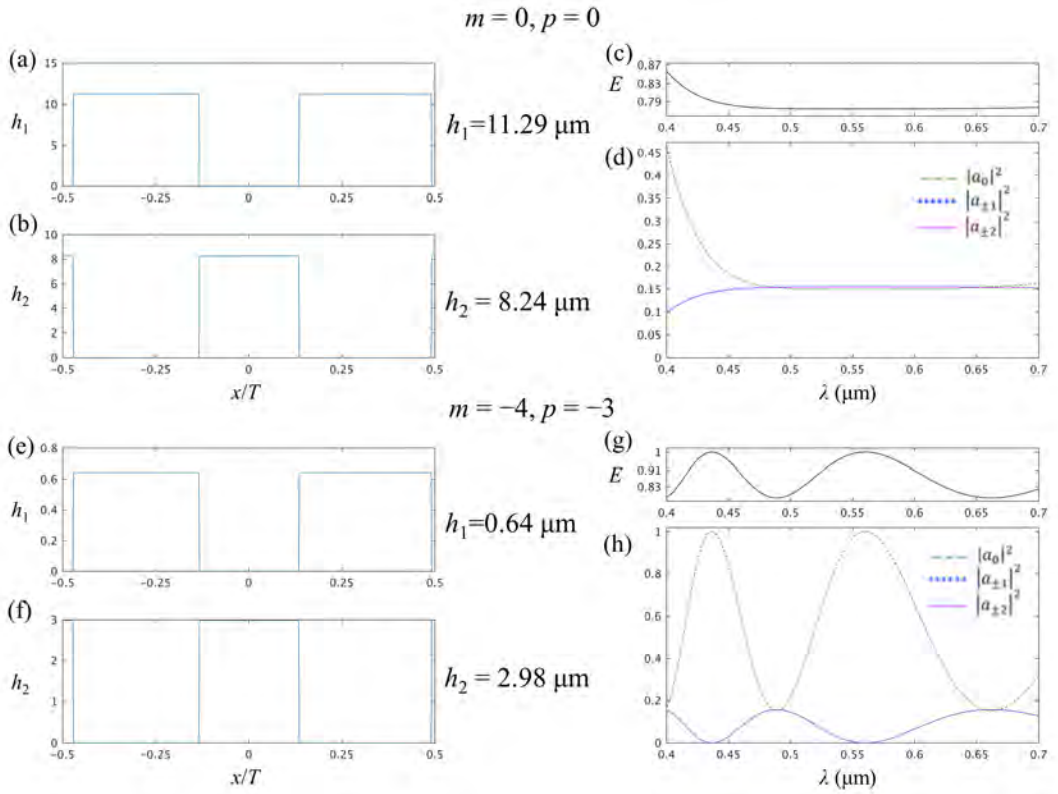

Fig. 5.11 (a,e) Height profile of grating layer 1. (b,f) Height profile for grating layer 2. (c,g) Total efficiency for grating as a function of wavelength. (d,h) Efficiency of each diffraction order. (a)-(d) m=0, p=0 solution (e)-(h) m = −4, p=−3 solution

Figure 5.12 shows the solutions when there is more than a 2π phase difference between λ_1 and λ_2. Figure 5.12(a) and 5.12(b) shows the height profiles of the grating layers when m=7 and p=5. Note that the summed height is still less than the m=0, p=0 solution. Once again there are oscillations in the efficiency as shown in Fig. 5.12(c). Figure 5.12(d) shows the efficiencies of the first five diffraction orders, and note there is now another wavelength that performs as designed in between λ = 486 nm and λ = 656 nm. This occurs because at some wavelength between λ = 486 nm and λ = 656 nm,

the phase is $\phi_0+2\pi$, which is equivalent to a phase of ϕ_0 for the transmitted electric field. If $m \neq p$, there are going to be $|m-p|-1$ points in between λ_1 and λ_2, where the transmitted phase is equivalent to a phase of ϕ_0. Figure 5.12(e) and 5.12(f) shows the height profiles for a solution when m = 21 and p=15, so there is a 12π phase difference between λ_1 and λ_2. Figure 5.12(g) and 5.12(h) shows there are 5 wavelengths in between λ = 486 nm and λ = 656 nm, where this grating has the equivalent performance to a grating with a phase of ϕ_0. As the phase difference between m and p grows, the performance becomes much more sensitive to wavelength changes. Therefore, adding an integer multiple of 2π to the phase between λ_1 and λ_2 has very poor performance for a broadband light source, however, this is an excellent design method to reduce the depth of the grating layers if the light source being used consists of two distinct wavelengths, like two lasers.

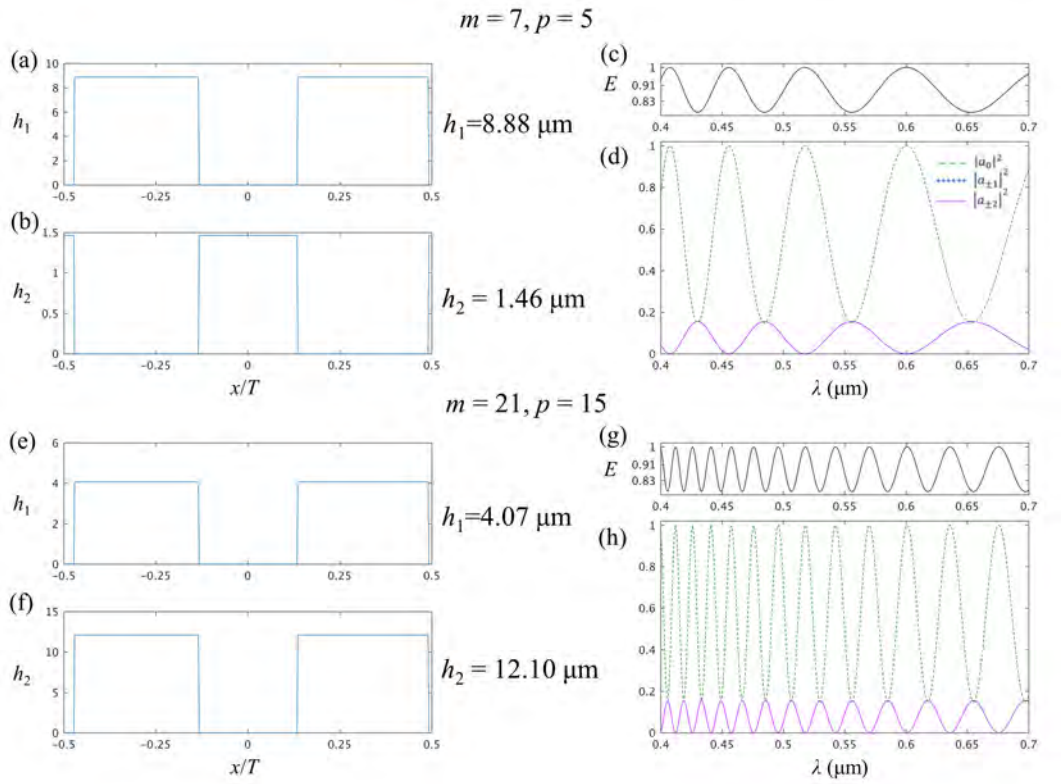

Fig. 5.12 (a,e) Height profile of grating layer 1. (b,f) Height profile for grating layer 2. (c,g) Total efficiency for grating. (d,h) Efficiency of each diffraction order. (a)-(d) m=7, p=5 solution (e)-(h) m = 21, p=15 solution

5.6 Conclusion

Simulations for an achromatic continuous phase grating doublet and an achromatic Dammann grating doublet are shown in this chapter. Sensitivities of achromatic beam-splitting grating doublets to lateral shifts between grating layers are studied in detail. It is shown that continuous phase grating doublets are extremely sensitive to lateral shifts between the grating layers, which makes fabrication extremely difficult. Grating profiles with a constant first spatial derivative decrease the performance change caused by lateral shifts of the grating layers. Grating profiles with a nearly constant first spatial derivative over a full period include blazed gratings, Dammann gratings, and multi-level phase gratings.

An achromatic Dammann grating doublet and an achromatic continuous phase grating doublet that split an incident beam into 5 equal energy outgoing beams at $\lambda_1 = 486$ nm and $\lambda_2 = 656$ nm are designed. A simulation of an achromatic Dammann grating doublet with a lateral shift between grating layers of one-hundredth of the period has 17 times better performance than the achromatic continuous phase grating doublet with the same lateral shift. Grating profiles with a constant first spatial derivative are superior designs for achromatic grating doublets, since they are significantly more resistant to lateral shifts between grating layers.

By studying the sensitivities to fabrication errors in the height, it is shown that certain height error combinations yield no performance change for a single wavelength. Additionally, for binary gratings it shown that if the height fabrication errors between the two grating layers combine to produce a phase error that is an integer multiple of 2π, there is no change in the far-field pattern at a particular wavelength. When averaging

over a large wavelength range, the performance of achromatic grating doublets can be preserved even with significant fabrication errors in the height of each grating layer. It is shown that the performance of an achromatic Dammann grating doublet can be improved by a factor of 215 if the height of the grating layers is chosen in order to minimize the performance change in the presence of height fabrication errors.

In addition to lateral shifts between grating layers and errors in the fabricated height of each layer, a grating layer may also be rotated or tilted with respect to the other grating layer. Both of these alignment errors further degrade the performance of a grating doublet. The diffraction pattern from a grating doublet is the convolution of the diffraction pattern from grating layer 1 with the diffraction pattern from grating layer 2. A rotation between grating layers rotates the diffraction patterns from each layer with respect to each other, which may degrade the performance of the grating doublet. A tilt in one grating layer changes the optical path length through that grating layer, which may further degrade the performance of the grating doublet.

One major disadvantage of achromatic grating doublets is the large height of the grating layers, which makes fabrication difficult. By adding an integer multiple of 2π between the designed phase at λ_1 and λ_2, the height of the grating layers can be reduced significantly. However, the performance of the achromatic grating between λ_1 and λ_2 degrades significantly, since the transmitted phase through the grating passes through an integer multiple of 2π radians. Therefore, the reduced depth grating cannot be used for a broadband light source with good performance. If the light source is two distinct wavelengths, like two different lasers, then adding an integer multiple of 2π phase

difference between λ_1 and λ_2 is an excellent way to design an achromatic grating with reduced grating layer depths.

6. ACHROMATIC PHASE GRATING FOR SCALE-INVARIANT DIFFRACTION PATTERNS

6.1 Introduction

Phase gratings are typically designed for high diffraction efficiency into one or multiple diffraction orders at a single wavelength. If the incident wavelength deviates from the designed wavelength, the diffraction efficiency changes. By using multiple layers, a diffractive optical element (DOE) can be achromatized with respect to diffraction efficiency for multiple wavelengths [80-82]. Instead of using multiple layers, the levels of single multi-level diffractive optical elements can be optimized to achromatize diffraction efficiency for specific multiple wavelengths [85,87]. By achromatizing a DOE using either of these two methods, diffraction efficiency performs as designed at two or more wavelengths. However, the diffraction angles associated with each diffraction order at these two wavelengths are not equal. Equal diffraction angles may be needed when coupling multiple wavelengths to a fiber or if multiple wavelength orders are overlapped at some image plane. Diffractive doublets have been designed for real imaging from one object point to one image point for two wavelengths [88]. However, the two wavelengths only overlap at a single image plane and do not have equal diffraction angles after the DOE. This chapter introduces a new type of grating called a scale-invariant achromatic (SIA) grating that has equal diffraction efficiency at two or more wavelengths and maintains equal diffraction angles for those wavelengths.

6.2 Theory

Assume that a lossless, one-dimensional grating is characterized by the periodic height function $h(x)$. In the absence of fabrication errors and using a thin grating approximation, the transmitted phase in air of a beam at position x is changed by an amount

$$\phi(x,\lambda) = h(x)2\pi\left(n(\lambda)-1\right)/\lambda, \tag{6.1}$$

where $n(\lambda)$ is the grating material refractive index at wavelength λ. A normally incident plane wave diffracts from a grating at an angle θ_m according to the grating equation

$$\sin(\theta_m) = m\lambda / T, \tag{6.2}$$

where m is the diffraction order of the grating and T is the period of the grating. Typically, the period of a grating is constant as a function of wavelength, thus small diffraction angles scale approximately linearly with λ. Suppose at λ_1 the period of the grating is T and the diffraction angle is θ_m. To maintain a diffraction angle of θ_m at λ_2, the grating period should be $T(\lambda_2/\lambda_1)$. Therefore, by changing the period as a function of wavelength, a grating can be designed to diffract light at the same angle for two different wavelengths.

For equal diffraction efficiencies at two different wavelengths, two grating layers with different refractive indices are placed in series, as shown in Fig. 6.1. Assume the desired phase profile at wavelength λ_1 is known and is represented by $\phi_1(x)$. The desired phase profile at wavelength λ_2 is also known and is represented by $\phi_2(x)$. The spatially varying optical path length (*OPL*) at these two wavelengths is given by

$$OPL(x,\lambda_1) = \left(n_1(\lambda_1)-1\right)h_1(x) + \left(n_2(\lambda_1)-1\right)h_2(x) = \frac{\phi_1(x)}{2\pi}\lambda_1 \tag{6.3}$$

$$OPL(x,\lambda_2) = (n_1(\lambda_2) - 1)h_1(x) + (n_2(\lambda_2) - 1)h_2(x) = \frac{\phi_2(x)}{2\pi}\lambda_2 \tag{6.4}$$

where n_1 and n_2 are the refractive indices for two different materials. Note that the sum of H_1 and H_2 are neglected in the OPL calculation, since they yield a constant OPD offset, which doesn't change the performance of the grating. Using this equation, the two unknown height profiles $h_1(x)$ and $h_2(x)$ are calculated. This method can be expanded to optimize gratings for more than two wavelengths by adding more layers. Achromatic grating doublets are designed by setting $\phi_1(x)$ equal to $\phi_2(x)$. In general, the phase profiles $\phi_1(x)$ and $\phi_2(x)$ at wavelength λ_1 and λ_2 are not equal or even similar. For example, height profiles can be calculated to create a blazed grating at λ_1 and a beam-splitting grating at λ_2.

For equal diffraction angles at λ_1 and λ_2, the periods should be T for $\phi_1(x)$ and $T(\lambda_2/\lambda_1)$ for $\phi_2(x)$. This chapter demonstrates SIA gratings by simulating an SIA blazed grating, a SIA radial grating, and a SIA beam-splitting grating with different beam-splitting properties at two wavelengths.

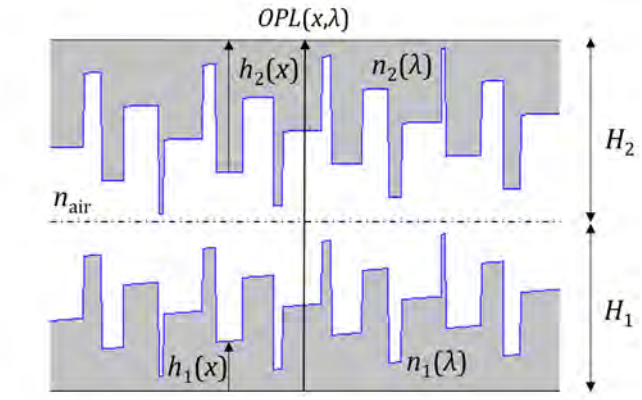

Fig. 6.1 OPL diagram for SIA grating. A SIA grating consists of two layers with different indices of refraction. Light is incident from below.

6.3 Scale-Invariant Achromatic Blazed Grating

This section simulates an achromatic blazed grating with equal diffraction angles at wavelengths λ_1=450 nm and λ_2=650 nm. Grating layer 1 is made from BK7 glass, and grating layer 2 is made from SF4 glass. At these wavelengths n_1(450 nm) = 1.525, n_1(650 nm) = 1.515, n_2(450 nm) = 1.786 ,and n_2(650 nm) = 1.748. This grating is designed to have a diffraction angle θ_m = 0.01 radians for λ_1 and λ_2 at m = +1. The period of the desired transmitted phase profile through both layers $\phi(x, \lambda)$ is designed to be T=45 μm at λ_1 and T=65 μm at λ_2. In Eq. (6.3) $\phi_1(x)$ is the phase profile of a blazed grating with T=45 μm. $\phi_2(x)$ is written as

$$\phi_2(x) = \phi_1\left(x\left(\lambda_1 / \lambda_2\right)\right) \tag{6.5}$$

to create a blazed grating with period T=65 μm,

The two unknown height profiles $h_1(x)$ and $h_2(x)$ for the SIA blazed grating are calculated using Eqs. (6.3) and (6.4). The height profiles for the two grating layers are shown in Fig. 6.2(a) and 6.2(b). The peak-to-valley heights are 67.5 μm and 45.6 μm. The height profile for each layer is a multi-level grating, in which each profile has a constant slope except at the transitions between levels. Note that the large heights may make these profiles difficult to fabricate.

The transmitted phase profile $\phi(x, \lambda)$ is shown for several wavelengths in Fig. 6.3. At λ=450 nm, $\phi(x, \lambda)$ is the designed blazed grating profile with a period of 45 μm. At λ=450 nm, 100% of the light is diffracted at θ_{+1}=0.01 radians. Likewise, at λ=650 nm $\phi(x, \lambda)$ is the designed blazed grating profile with a period of 65 μm. Thus, 100% of the light is diffracted at θ_{+1}=0.01 radians. At all other wavelengths, the grating deviates from the designed blazed grating profile. At λ=400 nm and λ=750 nm, the transmitted phase

$\phi(x, \lambda)$ deviates slightly from a perfect blazed grating, which leads to an efficiency slightly less than unity. At these wavelengths, the diffraction angle is not θ=0.01 radians. At some wavelengths, such as λ=500 nm, the performance is very poor. As shown in Fig. 6.3, the diffraction pattern consists of multiple low efficiency beams and $\phi(x, \lambda)$ deviates significantly from the transmitted phase for a blazed grating.

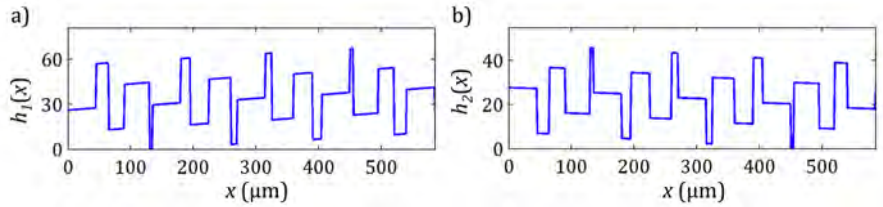

Fig. 6.2 Height profiles in microns (a) $h_1(x)$ and (b) $h_2(x)$ for SIA blazed grating. Peak-to-valley heights are 67.5 μm and 45.6 μm for $h_1(x)$ and $h_2(x)$, respectively.

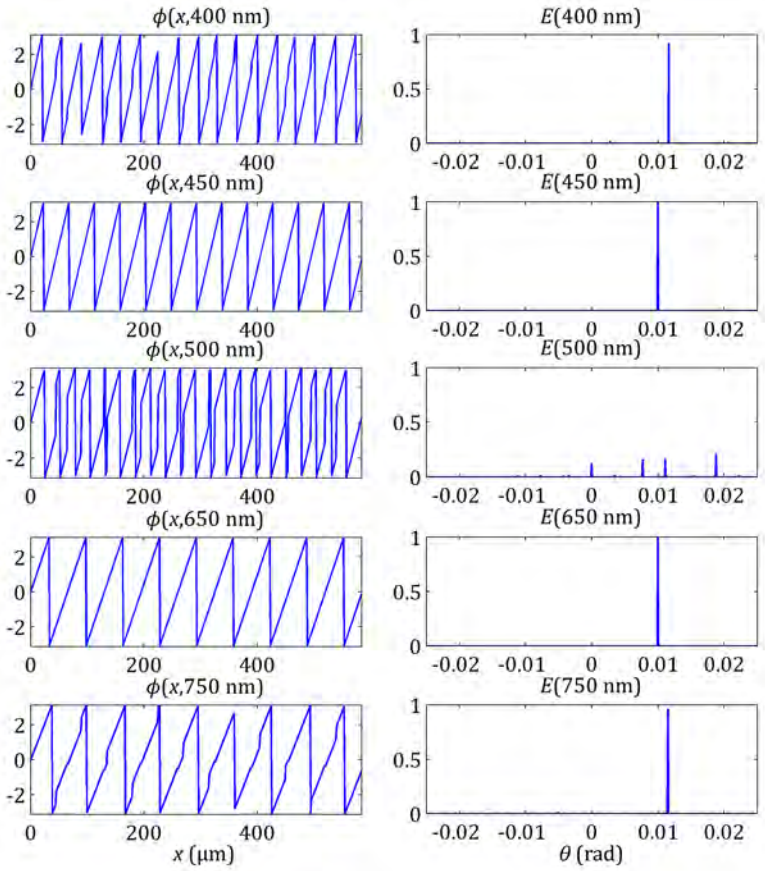

Fig. 6.3 (Left column) Transmitted phase $\phi(x)$ in radians through SIA blazed grating at λ=400 nm, λ=450 nm, λ=500 nm, λ=650 nm, and λ=750 nm. (Right column) Diffraction efficiency as a function of diffraction angle θ for SIA grating at indicated wavelengths.

Figure 6.4(a) shows the diffraction efficiency E of the diffraction order with the maximum diffraction efficiency. As designed, the diffraction efficiency is unity at λ_1=450 nm and λ_2=650 nm. As shown in Fig. 6.4(a), at some wavelengths, the efficiency drops significantly below unity. This drop is in stark contrast to an achromatic grating doublet, which has an efficiency of unity at two wavelengths, but different diffraction angles, which is shown in Fig. 6.4(b) [80]. For achromatic blazed grating doublets, the efficiency is unity at λ_1=450 nm and λ_2=650 nm, and is close to unity between λ_1 and λ_2. Figure 6.4(c) shows the diffraction angle θ of the diffraction order with maximum diffraction efficiency for the SIA blazed grating. At λ=450 nm and λ=650 nm, the diffraction angle is θ=0.01 radians. Figure 6.4(d) shows the diffraction angle θ of the diffraction order with maximum diffraction efficiency for an achromatic blazed grating doublet. The achromatic blazed grating doublet has a diffraction angle that increases approximately linearly with wavelength. Both the achromatic blazed grating doublet and the SIA blazed grating have designed diffraction efficiency at λ_1 and λ_2. While the achromatic grating doublet has superior diffraction efficiency between λ_1 and λ_2, the SIA grating has the advantage of equal diffraction angles at λ_1 and λ_2.

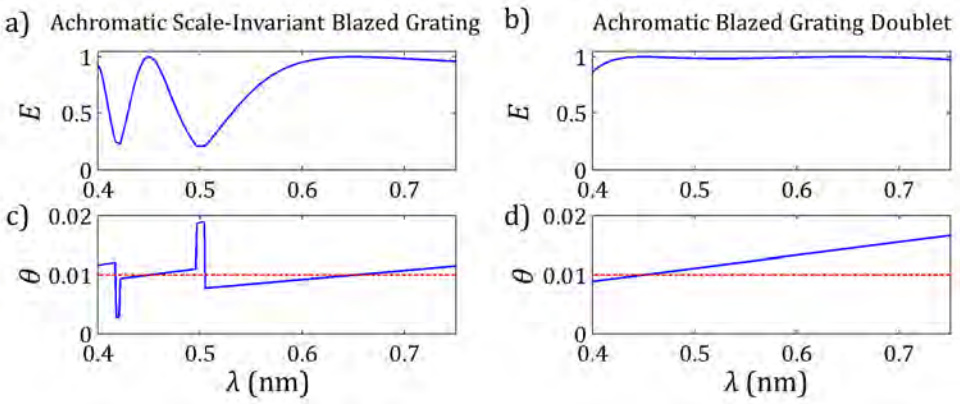

Fig. 6.4 Diffraction efficiency E of the diffraction order with maximum E as a function of wavelength for (a) SIA blazed grating and (b) achromatic blazed grating doublet. Diffraction angle θ for diffraction order with maximum diffraction efficiency as a function of wavelength for (c) SIA blazed grating and (d) achromatic blazed grating doublet.

6.4 Scale-Invariant Achromatic Radial Grating

This method is not limited to one-dimensional gratings. A radial SIA blazed grating is demonstrated that creates two ring diffraction patterns with the same radial diffraction angle θ_r=0.01 radians at wavelengths λ_1=480 nm and λ_2=620 nm. Grating layer 1 is made from BK7 glass, and grating layer 2 is made from SF4 glass. At these wavelengths n_1(480 nm) = 1.523, n_1(620 nm) = 1.516, n_2(480 nm) = 1.776, and n_2(620 nm) =1.751. The radial height profiles are shown in Fig. 6.5. The peak-to-valley heights are 108.07 μm and 73.36 μm.

The radial diffraction angle θ_r for the diffraction order with maximum diffraction efficiency as a function of wavelength is shown in Fig. 6.6(a). The diffraction patterns from the SIA radial grating at the design wavelengths are shown in Fig. 6.6(b) and 6.6(c). The phase profile at λ_1=480 nm is a radial grating with a period T=48 μm, so all light is sent into one ring with a radial diffraction angle θ_r=0.01 radians. At λ_2=620 nm the phase profile is a radial grating with a period T=62 μm, so all light is sent into one ring with a radial diffraction angle θ_r =0.01 radians. As shown in Fig. 6.6(a), the maximum diffraction efficiency is typically in the 0^{th} radial diffraction order when not at the designed wavelengths. For example, at λ=500 nm nearly all of the light is sent into the 0^{th} radial diffraction order, so no rings with significant energy are formed. At λ=600 nm and λ=660 nm, some light is sent into an outer ring, but the majority of the light is sent into the 0^{th} radial diffraction order. At λ=420 nm, there are multiple low amplitude ring patterns and none have the desired diffraction angle.

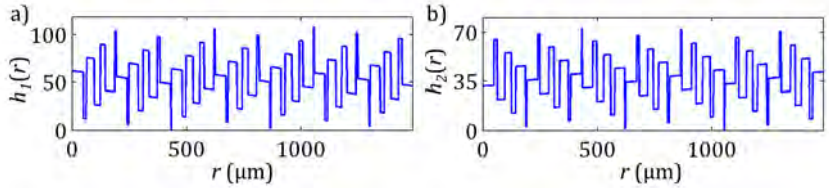

Fig. 6.5 Radial height profiles in microns (a) $h_1(r)$ and (b) $h_2(r)$ for SIA radial blazed grating. Peak-to-valley heights are 108.07 μm and 73.36 μm for $h_1(r)$ and h_2(r), respectively.

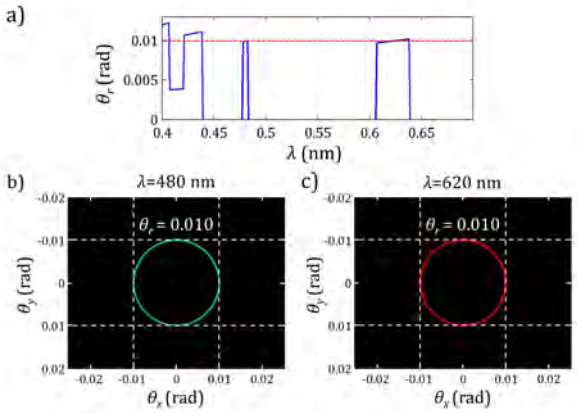

Fig. 6.6 (a) Radial diffraction angle as a function of wavelength. (b) Diffraction pattern of SIA radial blazed grating at design wavelengths λ=480 nm and λ=620 nm. (Diffracted light is psuedocolored to match the wavelength).

6.5 Scale-Invariant Achromatic Beam-Splitting Grating

It is not required in Eqs. (6.3) and (6.4) that $\phi_1(x)$ and $\phi_2(x)$ have similar desired phase profiles. This section demonstrates a SIA grating in which $\phi_1(x)$ is a binary phase Dammann grating that splits an incident beam into 3 equal diffraction efficiency beams at λ_1=450 nm, and $\phi_2(x)$ is a binary phase Dammann grating that splits an incident beam into 7 equal diffraction efficiency beams at λ_2=650 nm. The angular spacing between each successive diffraction order at these wavelengths is θ=0.01 radians. $\phi_1(x)$ has transitions between the high and low phases at x=(−0.25,0.25)T, where T=45 μm, with a maximum phase of 2.008 radians [62]. $\phi_2(x)$ has transitions between the high and low phases at x=(−0.430, −0.215, 0.215, 0.439)T, where T=65 μm, with a maximum phase of 2.473 radians [62].

The height profiles for this SIA grating are shown in Fig. 6.7. The peak-to-valley heights are 26.99 μm and 18.23 μm for $h_1(x)$ and $h_2(x)$, respectively. As shown in Fig. 6.8, the transmitted phase at λ=450 nm is the phase profile for a 1×3 beam-splitting grating with a period T=45 μm. At λ=650 nm the phase profile is the profile for a 1×7 beam-splitting grating with a period T=65 μm. At each of these wavelengths, the spacing between diffraction orders is θ=0.01. Once again, the performance when the wavelength is not λ=450 nm or λ=650 nm is poor, as evidenced by the diffraction pattern at λ=400 nm and λ=500 nm. At λ=400 nm the diffraction pattern contains several low efficiency diffraction orders that are not equal. At λ=500 nm nearly half the incident light is in the 0^{th} diffraction order. At 750 nm light is diffracted into 7 beams, but with unequal diffraction efficiency.

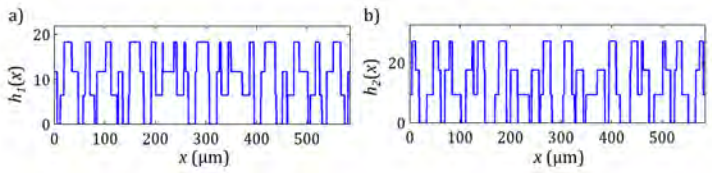

Fig. 6.7 Height profiles (a) $h_1(x)$ and (b) $h_2(x)$ for SIA beam-splitting grating that splits an incident beam into 3 beams at λ_1=450 nm and splits an incident beam into 7 beams at λ_2=650 nm. Peak-to-valley heights are 26.99 μm and 18.23 μm for $h_1(x)$ and $h_2(x)$, respectively.

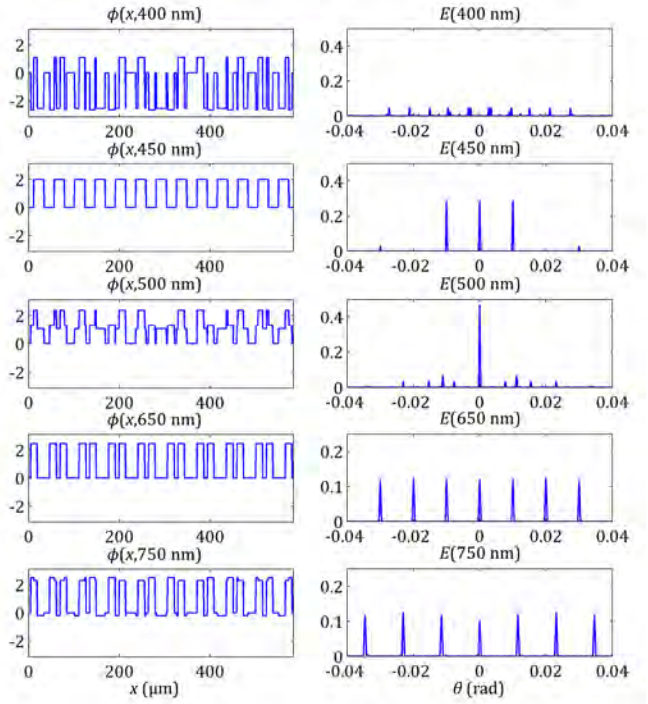

Fig. 6.8 (Left column) Transmitted phase $\phi(x)$ at λ=400 nm, λ=450 nm, λ=500 nm, λ=650 nm, and λ=750 nm through SIA grating that splits an incident beam into 3 beams with equal diffraction efficiency at λ_1=450 nm and splits an incident beam into 7 beams with equal diffraction efficiency at λ_2=650 nm. (Right column) Diffraction efficiency E as a function of diffraction angle θ for SIA grating at indicated wavelengths.

6.6 Reduced Height Scale-Invariant Achromatic Grating

This section explores if the method used in Sec. 5.5 to reduce the height of the achromatic grating can also be used reduce the height of SIA binary gratings by adding a integer multiple of 2π to the phase. For binary gratings, the optical path length at the two designed wavelengths is

$$OPL\left(x,\lambda_1\right)=\left(n_1\left(\lambda_1\right)-1\right)h_1(x)+\left(n_2\left(\lambda_1\right)-1\right)h_2(x)=\frac{b_1(x)\left(\phi_{10}+2\pi k\right)}{2\pi}\lambda_1 \qquad (6.6)$$

$$OPL\left(x,\lambda_2\right)=\left(n_1\left(\lambda_2\right)-1\right)h_1(x)+\left(n_2\left(\lambda_2\right)-1\right)h_2(x)=\frac{b_2(x)\left(\phi_{20}+2\pi p\right)}{2\pi}\lambda_2 \qquad (6.7)$$

where $b_i(x)$ is the base binary profile of the designed phase, ϕ_{i0} is the designed phase, and k and p are integers. For binary profiles, a phase of 2π can be added without any change in the designed performance. Now suppose an achromatic scale-invariant 1×3 Dammann

beam-splitting grating is achromatized at λ_1=450 nm and λ_2=650 nm with a diffraction angle of 0.01 radians at these two wavelengths. Figures 6.9(a) and 6.9(b) show the designed phase at λ_1=450 nm and λ_2=650 nm when k=0 and p=0 for this SIA grating. The 1×3 Dammann grating is a square wave with a 50% duty cycle and has a designed phase ϕ_{i0}=2.008 radians. Grating layer 1 is made from BK7 and grating layer 2 is made from SF4. At λ_1=450 nm the transmitted phase has a period of 45 μm, and at λ_2=650 nm the transmitted phase has a period of 65 μm.

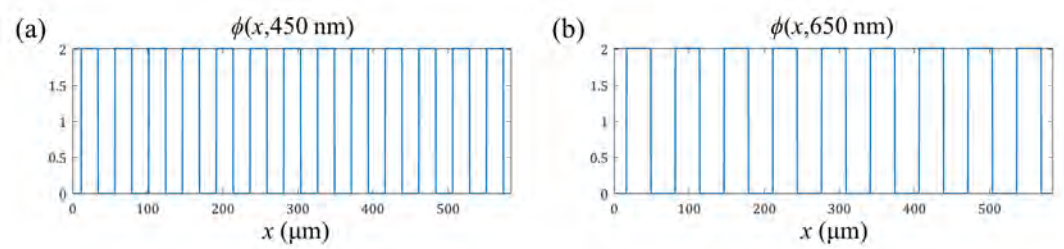

Fig. 6.9 (a) Designed phase for 1×3 Dammann SIA grating at 450 nm. (b) Designed phase for 1×3 Dammann SIA grating at 650 nm.

The heights of both grating layers are solved using Eqs. (6.6) and (6.7) for k=[-25,25] and p=[-25,25]. The logarithm of the sum of the two heights is shown in Fig. 6.10(a). The solution with a minimum summed height of the two layers is shown in Fig. 6.10(b). The minimum summed height occurs when k=0 and p=0. Note that there are 4 distinct levels in both of the grating layers. The structure of this profile is described in more detail later.

It is instructive to compare the heights of the SIA gratings to the heights of the achromatic grating doublets described in Chapter 5 for different k and p. For the achromatic grating doublet, $b_1(x) = b_2(x)$. For this example, the achromatic grating doublet has the transmitted phase shown in Fig. 6.9(a) at λ_1=450 nm and λ_2=650 nm. Figure 6.10(c) shows the minimum summed height as a function of k and p for the achromatic grating doublet, and Fig. 6.10(d) shows the solution with the minimum summed height.

As shown in Fig. 6.10, there are significant differences between the SIA grating and the achromatic grating doublet. Figure 6.10(a) shows the summed heights of the SIA gratings increases in a diamond-shaped pattern as $|k|$ and $|p|$ increase. However, Fig. 6.10(c) shows the minimum summed heights for the achromatic grating doublets are along a line. Furthermore, the grating profile for the SIA grating has 4 distinct levels, whereas the achromatic grating doublet only has 2. The following section attempts to explain these differences by analytically solving Eqs. (6.6) and (6.7).

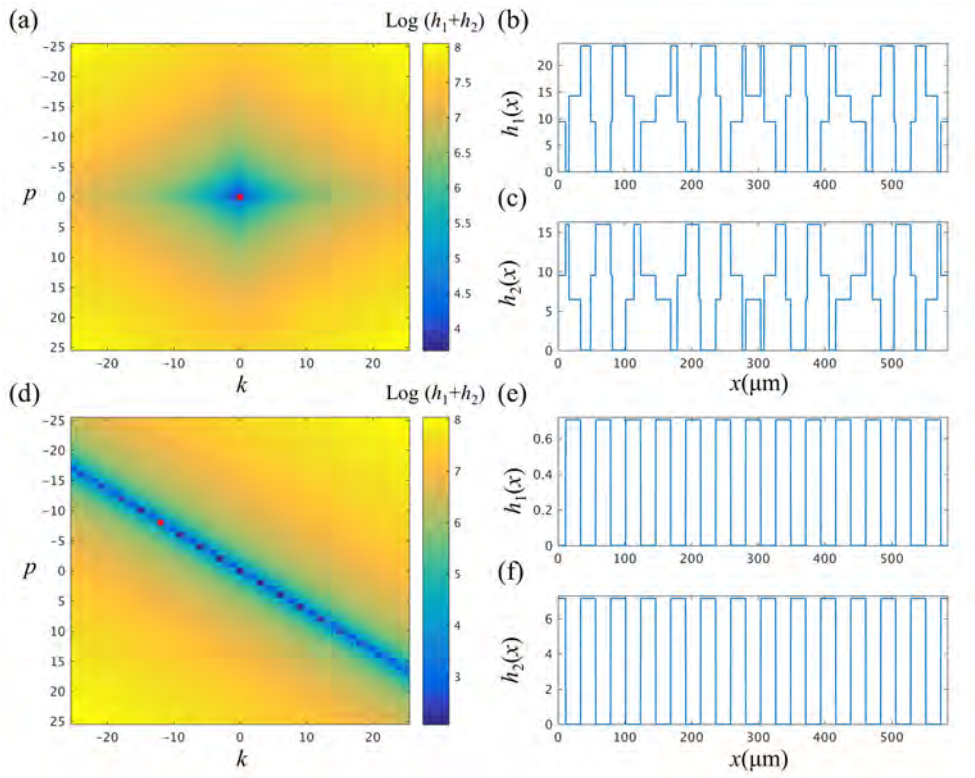

Fig. 6.10 (a) Log of summed heights for 1×3 SIA Dammann grating. (b) Height profiles for solution with minimum summed height. (c) Log of summed heights for 1×3 achromatic Dammann grating doublet. (d) Height profiles for solution with minimum summed height.

6.6.1 Theory

Solving Eqs. (6.6) and (6.7) analytically yields the following equations for $h_1(x)$ and $h_2(x)$:

$$h_1(x)=\frac{b_1(x)\left(n_2\left(\lambda_2\right)-1\right)\lambda_1\left(2\pi k+\phi_{10}\right)-b_2(x)\left(n_1\left(\lambda_2\right)-1\right)\lambda_2\left(2\pi p+\phi_{20}\right)}{2\pi\left(n_1\left(\lambda_2\right)+n_2\left(\lambda_1\right)-n_1\left(\lambda_2\right)n_2\left(\lambda_1\right)+n_1\left(\lambda_1\right)n_2\left(\lambda_2\right)-n_1\left(\lambda_1\right)-n_2\left(\lambda_2\right)\right)}, \quad (6.8)$$

$$h_2(x)=\frac{-b_1(x)\left(n_2\left(\lambda_1\right)-1\right)\lambda_1\left(2\pi k+\phi_{10}\right)+b_2(x)\left(n_1\left(\lambda_1\right)-1\right)\lambda_2\left(2\pi p+\phi_{20}\right)}{2\pi\left(n_1\left(\lambda_2\right)+n_2\left(\lambda_1\right)-n_1\left(\lambda_2\right)n_2\left(\lambda_1\right)+n_1\left(\lambda_1\right)n_2\left(\lambda_2\right)-n_1\left(\lambda_1\right)-n_2\left(\lambda_2\right)\right)}. \quad (6.9)$$

The following definitions are made, so these equations can be written more succinctly.

$$\Delta n_{ij}=n_i(\lambda_j)-1. \quad (6.10)$$

$$D=2\pi\left(n_1\left(\lambda_2\right)+n_2\left(\lambda_1\right)-n_1\left(\lambda_2\right)n_2\left(\lambda_1\right)+n_1\left(\lambda_1\right)n_2\left(\lambda_2\right)-n_1\left(\lambda_1\right)-n_2\left(\lambda_2\right)\right). \quad (6.11)$$

Substituting Eqs. (6.10) and (6.11) into Eqs. (6.8) and (6.9) yields

$$h_1(x)=\frac{b_1(x)\Delta n_{22}\lambda_1\left(2\pi k+\phi_{10}\right)-b_2(x)\Delta n_{12}\lambda_2\left(2\pi p+\phi_{20}\right)}{D}, \quad (6.12)$$

$$h_2(x)=\frac{-b_1(x)\Delta n_{21}\lambda_1\left(2\pi k+\phi_{10}\right)+b_2(x)\Delta n_{11}\lambda_2\left(2\pi p+\phi_{20}\right)}{D}. \quad (6.13)$$

Since $b_1(x)$ and $b_2(x)$ are either 0 or 1, there are 4 possible heights for $h_1(x)$ and $h_2(x)$ in general. The four possible values for $h_1(x)$ and $h_2(x)$ are written as vectors $\boldsymbol{H}_1$ and $\boldsymbol{H}_2$:

$$\boldsymbol{H}_1=\begin{Bmatrix}H_{11}\\H_{12}\\H_{13}\\H_{14}\end{Bmatrix}=\begin{Bmatrix}\dfrac{\Delta n_{22}\lambda_1\left(2\pi k+\phi_{10}\right)-\Delta n_{12}\lambda_2\left(2\pi p+\phi_{20}\right)}{D}\\[2ex]\dfrac{\Delta n_{22}\lambda_1\left(2\pi k+\phi_{10}\right)}{D}\\[2ex]\dfrac{-\Delta n_{12}\lambda_2\left(2\pi p+\phi_{20}\right)}{D}\\[2ex]0\end{Bmatrix} \quad (6.14)$$

$$\boldsymbol{H}_2=\begin{Bmatrix}H_{21}\\H_{22}\\H_{23}\\H_{24}\end{Bmatrix}=\begin{Bmatrix}\dfrac{-\Delta n_{21}\lambda_1\left(2\pi k+\phi_{10}\right)+\Delta n_{11}\lambda_2\left(2\pi p+\phi_{20}\right)}{D}\\[2ex]\dfrac{-\Delta n_{21}\lambda_1\left(2\pi k+\phi_{10}\right)}{D}\\[2ex]\dfrac{\Delta n_{11}\lambda_2\left(2\pi p+\phi_{20}\right)}{D}\\[2ex]0\end{Bmatrix} \quad (6.15)$$

In general the values for $\boldsymbol{H}_1$ and $\boldsymbol{H}_2$ can be negative, however, when fabricating the grating layers, the layers need to be referenced from the minimum value. Therefore, the total heights fabricated are

$$\Delta H_1 = \max\{\boldsymbol{H}_1\} - \min\{\boldsymbol{H}_1\} \tag{6.16}$$

$$\Delta H_2 = \max\{\boldsymbol{H}_2\} - \min\{\boldsymbol{H}_2\} \tag{6.17}$$

Before analyzing why k and p doesn't decrease $h_1(x)+h_2(x)$ for the SIA gratings, it is instructive to examine how the height is reduced for achromatic grating doublets. Assume $b_1(x)=b_2(x)$, which occurs for the achromatic grating doublets. When $b_1(x)=b_2(x)$, there are now only two possible values for $\boldsymbol{H}_{\mathrm{i}}$, $H_{\mathrm{i}1}$ and $H_{\mathrm{i}2}$. The height of each grating layer is the difference between these two possible values, which is written as

$$\Delta H_1 = \left| \frac{\Delta n_{22}\lambda_1\left(2\pi k + \phi_{10}\right) - \Delta n_{12}\lambda_2\left(2\pi p + \phi_{20}\right)}{D} \right|, \tag{6.18}$$

$$\Delta H_2 = \left| \frac{-\Delta n_{21}\lambda_1\left(2\pi k + \phi_{10}\right) + \Delta n_{11}\lambda_2\left(2\pi p + \phi_{20}\right)}{D} \right|. \tag{6.19}$$

To determine the values for k and p that yield solutions less than k=0 and p=0, let us first define the height of grating layer 1 when k=0 and p=0 as

$$\left|\Delta H_{10}\right| = \left| \frac{\Delta n_{22}\lambda_1\phi_{10} - \Delta n_{12}\lambda_2\phi_{20}}{D} \right|. \tag{6.20}$$

The height grating layer 1 is less than the k=0 and p=0 solution when

$$\left| \frac{\Delta H_{10} + 2\pi\left(\Delta n_{22}\lambda_1 k - \Delta n_{12}\lambda_2 p\right)}{D} \right| < \left|\Delta H_{10}\right|, \tag{6.21}$$

which can be rewritten as

$$-\left|\Delta H_{10}\right| < \frac{\Delta H_{10} + 2\pi\left(\Delta n_{22}\lambda_1 k - \Delta n_{12}\lambda_2 p\right)}{D} < \left|\Delta H_{10}\right|. \tag{6.22}$$

After simplifying, the following inequality is

$$\frac{\left(-\left|\Delta H_{10}\right|-\Delta H_{10}\right)D}{2\pi}<\Delta n_{22}\lambda_1 k-\Delta n_{12}\lambda_2 p<\frac{\left(\left|\Delta H_{10}\right|-\Delta H_{10}\right)D}{2\pi} \tag{6.23}$$

Therefore, the heights less than $|\Delta H_{10}|$ are along the line with a slope of

$$\Delta H_{1slope}=\frac{\Delta n_{22}\lambda_1}{\Delta n_{12}\lambda_2} \tag{6.24}$$

Not all values of k and p on this line result in heights less than $|\Delta H_{10}|$, the inequality in Eq. (6.23) must also be satisfied. A similar analysis of h_2 results in the heights that are less than the k=0 and p=0 solution lying along the line with a slope of

$$\Delta H_{2slope}=\frac{\Delta n_{21}\lambda_1}{\Delta n_{11}\lambda_2} \tag{6.25}$$

Figure 6.11(a) and 6.11(b) shows the logarithm of ΔH_1 and ΔH_2 for the 1×3 Dammann achromatic grating doublet. The lines with slopes from Eqs. (6.24) and (6.25) are shown as red lines, which lies along the minimum values of ΔH_1 and ΔH_2. Since there are only two possible values in the grating profile, there exist several combinations of k and p that result in solutions with a height less than the k=0 and p=0 solution.

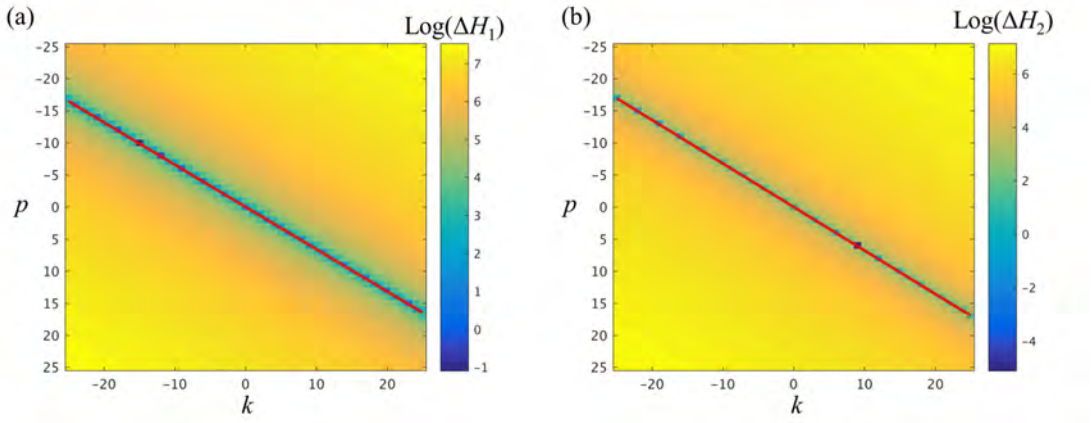

Fig. 6.11 (a) Log of ΔH_1 for 1×3 SIA Dammann grating. (b) Log of ΔH_2 for 1×3 SIA Dammann grating. Red line is the line with slopes given by Eqs. (6.24) and (6.25) .

Now let's examine why the heights of SIA gratings are not reduced for this 1×3 Dammann grating. Let's define the k=0 and p=0 height for grating layer 1 as

$$\Delta H_1 = \max\left\{\begin{array}{l}\dfrac{\Delta n_{22}\lambda_1\phi_{10} - \Delta n_{12}\lambda_2\phi_{20}}{D} \\ \dfrac{\Delta n_{22}\lambda_1\phi_{10}}{D} \\ \dfrac{-\Delta n_{12}\lambda_2\phi_{20}}{D} \\ 0\end{array}\right\} - \min\left\{\begin{array}{l}\dfrac{\Delta n_{22}\lambda_1\phi_{10} - \Delta n_{12}\lambda_2\phi_{20}}{D} \\ \dfrac{\Delta n_{22}\lambda_1\phi_{10}}{D} \\ \dfrac{-\Delta n_{12}\lambda_2\phi_{20}}{D} \\ 0\end{array}\right\}. \tag{6.26}$$

Let's assume the max value is H_{12} and the minimum value is H_{13}. Under these assumptions, the total height of grating layer 1 when k=0 and p=0 is

$$\Delta H_1 = \frac{\Delta n_{22}\lambda_1\phi_{10}}{D} + \frac{\Delta n_{12}\lambda_2\phi_{20}}{D}. \tag{6.27}$$

To maintain the assumption that the H_{12} and H_{13} are the max and min values when k and p are non-zero, then the following inequality must be true:

$$\frac{\Delta n_{22}\lambda_1\left(2\pi k + \phi_{10}\right)}{D} > \frac{-\Delta n_{12}\lambda_2\left(2\pi p + \phi_{20}\right)}{D}. \tag{6.28}$$

Solving for p yields

$$p > -\frac{\Delta n_{22}\lambda_1 2\pi k + \Delta n_{22}\lambda_1\phi_{10} + \Delta n_{12}\lambda_2\phi_{20}}{\Delta n_{12}\lambda_2 D 2\pi}. \tag{6.29}$$

To reduce the height ΔH_1, then H_{12} needs to decrease and H_{13} needs to increase, which yields the following inequalities:

$$\frac{\Delta n_{22}\lambda_1\left(2\pi k + \phi_{10}\right)}{D} < \frac{\Delta n_{22}\lambda_1\phi_{10}}{D}, \tag{6.30}$$

$$\frac{-\Delta n_{12}\lambda_2\left(2\pi p + \phi_{20}\right)}{D} > \frac{-\Delta n_{12}\lambda_2\phi_{20}}{D}. \tag{6.31}$$

Therefore, k<0 and p <0.

To maintain this assumption, H_{11} can't be greater than H_{12}. Therefore, the following inequality must also be true

$$\frac{\Delta n_{22}\lambda_1\left(2\pi k+\phi_{10}\right)-\Delta n_{12}\lambda_2\left(2\pi p+\phi_{20}\right)}{D}<\frac{\Delta n_{22}\lambda_1\left(2\pi k+\phi_{10}\right)}{D}, \tag{6.32}$$

which results in

$$p<-\frac{\phi_{20}}{2\pi}. \tag{6.33}$$

Additionally, H_{11} needs to be larger $H_{13:}$

$$\frac{\Delta n_{22}\lambda_1\left(2\pi k+\phi_{10}\right)-\Delta n_{12}\lambda_2\left(2\pi p+\phi_{20}\right)}{D}>\frac{-\Delta n_{12}\lambda_2\left(2\pi p+\phi_{20}\right)}{D}, \tag{6.34}$$

which results in

$$k>-\frac{\phi_{10}}{2\pi}. \tag{6.35}$$

The bounds for p to maintain this assumption are

$$-\frac{\Delta n_{22}\lambda_1 2\pi k+\Delta n_{22}\lambda_1\phi_{10}+\Delta n_{12}\lambda_2\phi_{20}}{\Delta n_{12}\lambda_2 D2\pi}<p<-\frac{\phi_{20}}{2\pi}, \tag{6.36}$$

which simplifies to

$$-\frac{\Delta n_{22}\lambda_1}{\Delta n_{12}\lambda_2 D}k-\frac{\Delta n_{22}\lambda_1\phi_{10}}{\Delta n_{12}\lambda_2 D2\pi}-\frac{\phi_{20}}{D2\pi}<p<-\frac{\phi_{20}}{2\pi}. \tag{6.37}$$

The bounds for k to maintain the p inequality is

$$-\frac{\Delta n_{22}\lambda_1}{\Delta n_{12}\lambda_2 D}k-\frac{\Delta n_{22}\lambda_1\phi_{10}}{\Delta n_{12}\lambda_2 D2\pi}-\frac{\phi_{20}}{D2\pi}<-\frac{\phi_{20}}{2\pi}, \tag{6.38}$$

which simplifies to

$$k>-\frac{\phi_{20}}{2\pi}\left(1-D\right)\frac{\Delta n_{12}\lambda_2}{\Delta n_{22}\lambda_1}-\frac{\phi_{10}}{2\pi}. \tag{6.39}$$

Therefore, the values of k and p that maintain the assumption are

$$-\frac{\phi_{20}}{2\pi}(1-D)\frac{\Delta n_{12}\lambda_2}{\Delta n_{22}\lambda_1}-\frac{\phi_{10}}{2\pi}<k<0 \tag{6.40}$$

$$-\frac{\Delta n_{22}\lambda_1}{\Delta n_{12}\lambda_2 D}k-\frac{\Delta n_{22}\lambda_1\phi_{10}}{\Delta n_{12}\lambda_2 D2\pi}-\frac{\phi_{20}}{D2\pi}<p<-\frac{\phi_{20}}{2\pi} \tag{6.41}$$

Note that k and p are integers, so there are limited values of k and p that can satisfy these two equations. Also note these are only the constraints for ΔH_1, there are a similar set of constraints for ΔH_2. Additionally, these are only the constraints to maintain the assumption that H_{12} is the max and H_{13} is the minimum. This is not the only possible configuration for the grating layer. For instance, H_{13} could flip to be the maximum and H_{12} could flip to be the minimum value. This configuration comes with its own set of constraints for k and p to maintain that assumption. Since there are 4 possible values for the heights of the grating layers it is much more difficult to find values for k and p that decrease the overall height of the grating layer. In order to decrease one element in H_1, it is often the case that another element grows in size, which is the reason the overall height of the grating tends to grow as $|k|>0$ and $|p|>0$. Note these inequalities prove that in general it is possible to find a set of k and p that reduces the overall height of the grating if the inequalities are satisfied, but it is not as likely as the achromatic grating doublet. There may exist a set of k and p that does reduce the overall height of the grating, but due to the constraints set by the 4 distinct levels, there are significantly fewer values of k and p that reduce the overall height than the achromatic grating doublet.

6.7 Conclusion

A method to achromatize a phase grating for designed diffraction efficiency at two wavelengths, while also diffracting at the same angle, is demonstrated by designing several examples. A blazed grating is designed to diffract all incident light into the +1

diffraction order at an angle of θ=0.01 radians at λ=450 nm and λ=650 nm. This method is shown to also work in two dimensions by demonstrating a SIA radial grating, which diffracts all incident light into rings, where the angle of the rings are θ_r=0.01 radians at both λ=480 nm and λ=620 nm. It is not required to have similar phase profiles at different wavelengths, as demonstrated by the SIA beam-splitting grating. This grating splits an incident beam into 3 beams with equal diffraction efficiency at λ=450 nm and 7 beams with equal diffraction efficiency at λ=650 nm. The spacing between diffraction orders at both wavelengths is θ=0.01 radians.

As long as the desired phase profile at each desired wavelength is known, a SIA grating can be designed. SIA gratings not only have designed diffraction efficiency at two wavelengths, but the diffraction angle is also equal at those two wavelengths. This method is easily expanded to more than 2 wavelengths by adding more layers into the grating. Furthermore, it is not necessary to constrain the diffraction angle to be equal at two wavelengths. The diffraction angle can also be designed to be any value at two wavelengths. For example, the diffraction angle for λ_2 can be designed to be three times the diffraction angle of λ_1.

A method is explored in an attempt to reduce the overall height of binary SIA gratings, since they are often quite large. This method adds a constant multiple of 2π to the designed phase when solving for the heights. In contrast to the significantly reduced heights shown in Chapter 5 using this method, binary SIA gratings are less likely to reduce the height using this method. This is due to the fact that there are 4 distinct levels in an SIA for binary phase profiles. While trying to minimize one of these 4 levels, the

other levels in the SIA grating tend to grow, which makes reducing the height using this method not possible for some binary phase profiles.

7. DESIGNS FOR MULTI-WAVELENGTH ARRAY CONFOCAL MICROSCOPES

7.1 Introduction

The designs for the array confocal microscopes presented in Chapter 2 and 4 are only designed for a single wavelength. However, multiple fluorophores are often used to stain a sample, since each fluorophore adds additional information about that sample being studied. The diffraction gratings used in Chapters 2 and 4 are optimized for a single wavelength, which means the performance degrades when the wavelength is not the designed wavelength. Additionally, the diffraction angle is different for different wavelengths. Therefore, it is not possible to couple the light from multiple wavelengths into a single fiber using gratings optimized for a single wavelength.

This chapter explores possible designs to excite and detect multiple fluorophores for the design of the array confocal microscope presented in Chapter 4. Since Chapters 5 and 6 explored designs for achromatic beam-splitting gratings, possible achromatic array confocal microscope designs using these gratings are conceptualized. For this chapter, it is assumed that the lenses used in the array confocal microscope are optimized for diffraction-limited performance throughout the entire optical system at multiple wavelengths. Therefore, the excitation spots at different wavelengths are focused onto the same position in the sample plane.

7.2 Excitation

The excitation path and the detection path in the array confocal microscope need to be modified for multiple wavelength excitation and detection. This section shows the modifications necessary in the excitation system for multi-wavelength functionality.

7.2.1 Two Beam-splitting Gratings

One method to transfer multiple excitation wavelengths into the confocal arm is to use two beam-splitting gratings to independently couple the light from two lasers. Figure 7.1 shows this concept. Each beam-splitting grating profile is optimized for a single wavelength, just like in Chapters 3 and 4. The diffracted light from the beam-splitting gratings is coupled into a single-mode fiber, as demonstrated in Chapter 4. Therefore, 18 single-mode fibers are required for this set-up.

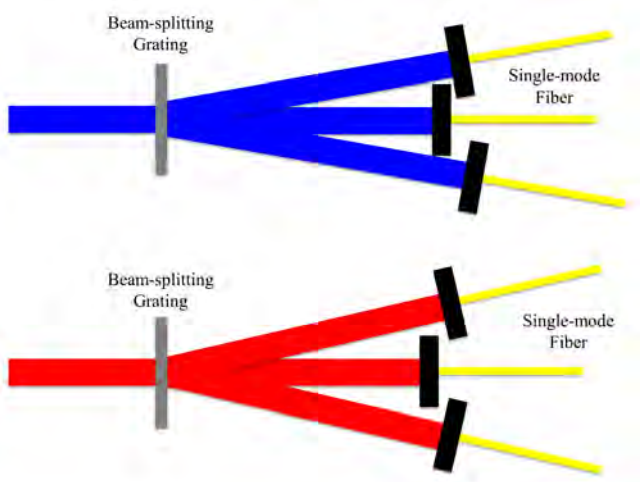

Fig. 7.1 Using two single-wavelength beam-splitting gratings to couple two multiple wavelengths of light into single-mode fibers.

The other end of each of these fibers is positioned on one of the confocal arms from Chapter 4. These two wavelengths should be collimated and collinear when incident on the galvo scanner. The collinearity is required to ensure each wavelength is focused at the same position in the sample plane. One possible design is shown in Fig. 7.2. On each

confocal arm, the two wavelengths are made collinear by reflecting off a dichroic mirror. In Chapter 4, a dichroic mirror sends the excitation beam to the galvo scanner, since the fluorescence is much higher in wavelength than the excitation beam. Therefore, a dichroic mirror is used to easily distinguish the two wavelength ranges. When using multiple excitation wavelengths, one excitation wavelength may overlap with the fluorescence wavelengths, so the dichroic mirror is replaced with a generalized beamsplitter. This could be a 90/10 beamsplitter that reflects 10% of the light and transmits 90% of the light, since the fluorescence is significantly weaker than the excitation light. The fluorescence from multiple fluorophores is transmitted through the beamsplitter where it is sent to the detection system. A detection system for multiple wavelengths is described in Sec. 7.3.

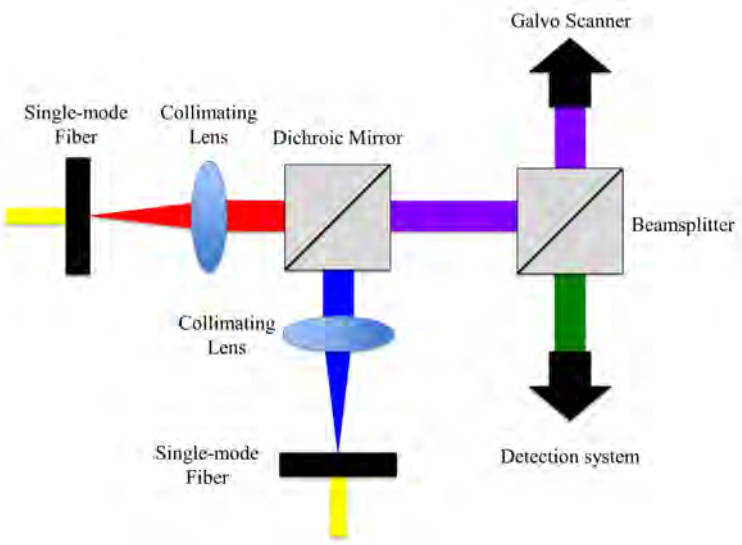

Fig. 7.2 Two wavelengths from a single-mode fiber are combined into a collinear beam after a dichroic mirror. Both wavelengths are reflected off a beamsplitter and are incident on a galvo scanner. The fluorescence from multiple wavelengths is transmitted through the beamsplitter and is detected.

7.2.2 Achromatic Beam-splitting Grating Doublet

Suppose instead of using two single wavelength beam-splitting gratings to couple the excitation light into the fibers, an achromatic beam-splitting grating doublet from

Chapter 5 is used. This achromatic grating doublet has the designed performance at two wavelengths, however the diffraction angle is different. Figure 7.3 shows a diagram to couple the light from this grating into single-mode fibers. Two lasers are combined using a dichroic mirror and are required to be collinear when incident on the grating. The grating splits the beam into *N* equal energy beams at two different wavelengths, however, as shown in Fig. 7.3 the diffraction angle is different. Therefore, 17 single-mode fibers need to be used to couple all beams into fibers, since the diffraction angle is equal for the zero-order diffracted beam for each wavelength. One complication arises since the diffraction angle between wavelengths is relatively small, which requires a very long distance between the grating doublet and the fiber to separately couple each wavelength without blocking the other wavelength. The light from the fibers is then attached to the confocal arm as shown in Fig. 7.2. Overall, using this grating doesn't yield many advantages for this set-up. Using two single wavelength beam-splitting gratings is most likely easier to set up and align than using an achromatic grating doublet.

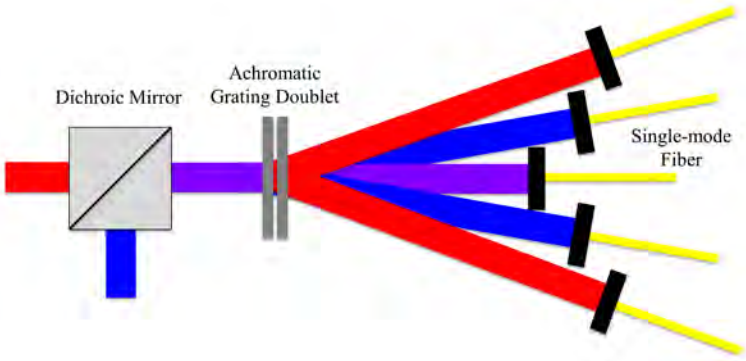

Fig. 7.3 Coupling multiple wavelengths into single-mode fibers using an achromatic grating doublet.

7.3.3 Scale-Invariant Achromatic Beam-splitting Grating

A scale-invariant achromatic beam-splitting grating, not only has the designed beam-splitting performance at two wavelengths, but it also has equal diffraction angles at two

wavelengths. Using this grating greatly simplifies the set-up for multiple wavelengths, as shown in Figure 7.4. A dichroic mirror is used to combine two lasers and make them collinear. The two wavelengths are split into *N* beams and then both coupled into the same single-mode fiber. For this case, the coupling lens for the single mode fiber should be optimized for no axial or lateral chromatic aberrations at these two wavelengths.

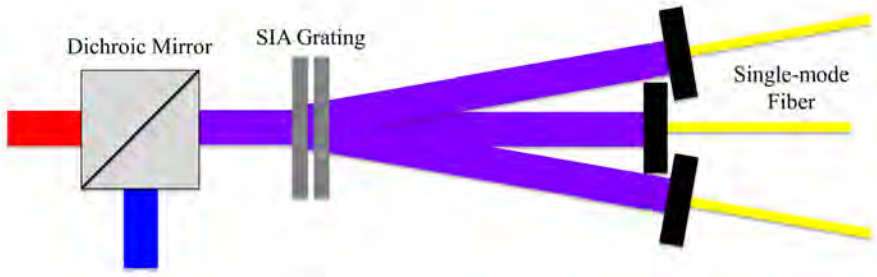

Fig. 7.4 Coupling multiple wavelengths into single-mode fibers using a scale-invariant achromatic grating doublet.

The design for the confocal arm is shown in Fig. 7.5, which is very similar to the excitation arm in Chapter 4. The dichroic mirror may be replaced with a beamsplitter, but otherwise the excitation arm is identical. The collimating lens should be achromatic. Since multiple fluorophores need to be separately detected, the detection arm needs to be modified.

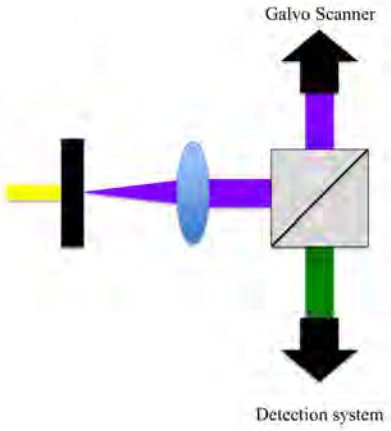

Fig. 7.5 Confocal arm for system when two wavelengths are coupled into one single-mode fiber.

7.2 Detection

Since multiple fluorophores are fluorescing in different wavelength bands, each of these fluorophores needs to be separately measured. The process of detecting multiple fluorophores can be accomplishing using two detection fibers or a single detection fiber in each channel.

7.2.1 Two Detection Fibers

One method is to use two detection fibers to detect the fluorescence from multiple fluorophores. A dichroic mirror is placed after the beamsplitter in the confocal arm and appropriate long pass filters are placed in each detection channel, as shown in Fig. 7.6. A focusing lens and fiber diameter can then be optimized for each wavelength separately. Each wavelength coupled to a fiber is then imaged and detected on its own PMT array. While this set-up allows for separate optimization and detection for each fluorophore, the confocal arm has the potential to be extremely large, which would make mounting it at a specified angle on a mechanical mount very difficult.

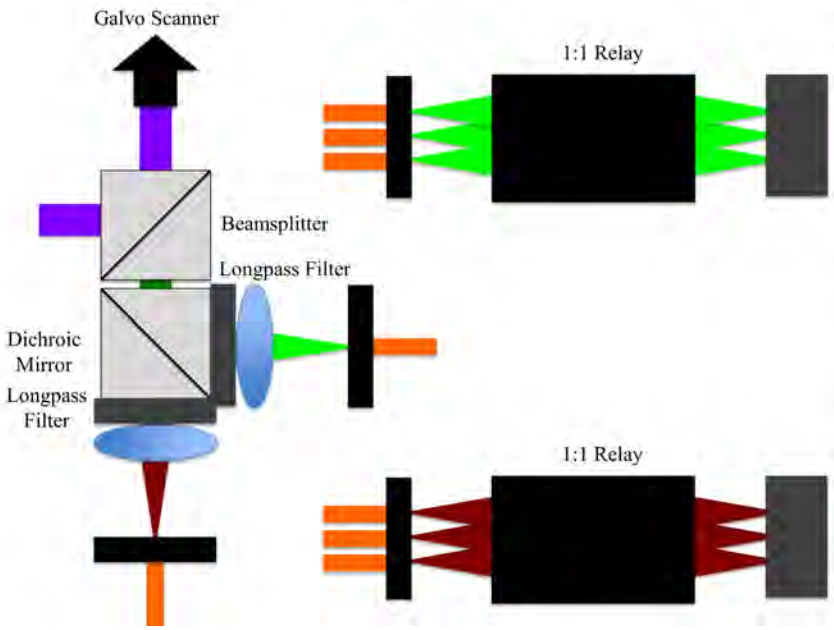

Fig. 7.6 Layout to detect two fluorophores using two detection fibers in the confocal arm. A 1:1 relay images the fluorescence from each fiber onto a channel of the PMT array.

7.2.2 Single Detection Fiber

Instead of using two detection fibers to detect multiple fluorophores, a single fiber can be used in the confocal arm to collect the fluorescence signal, as shown in Fig. 7.7. A notch filter is used to block the two excitation wavelengths and transmit the fluorescence signal. An achromatic focusing lens focuses the fluorescence onto the end of a multi-mode fiber. One caveat that arises with this set-up is the Airy disk radius is different at separate wavelengths. While the fiber radius can be optimized to be one Airy disk radius at the shorter wavelength, the fiber radius is going to be smaller than the Airy disk radius at the longer wavelength. Therefore, the confocal resolution and optical sectioning at the two wavelengths differs slightly.

After the fluorescence is collected by the fiber, the other end of the fiber is connected to a 1:1 relay system. A dichroic mirror is used to separate the fluorescence from the different fluorophores. A separate PMT array for each fluorophore is used to detect the fluorescence signal from multiple channels.

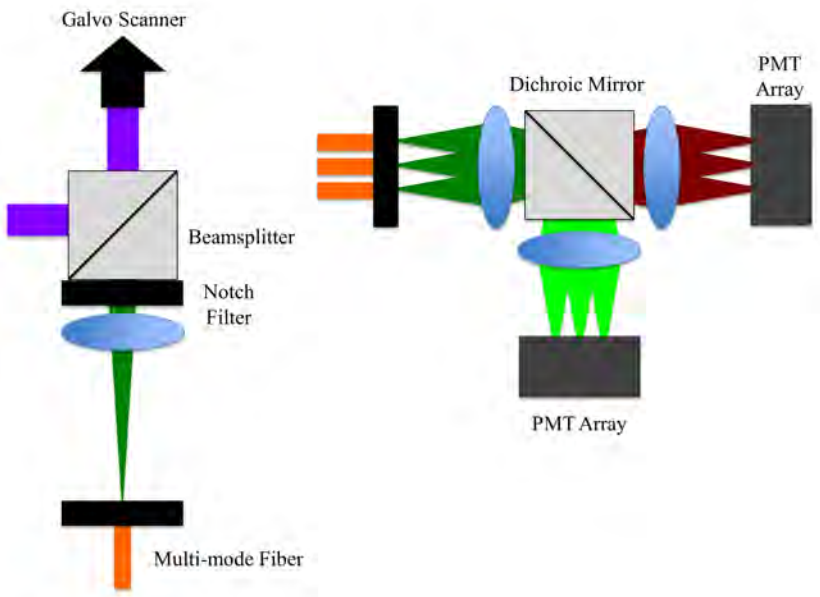

Fig. 7.7 Layout to detect two fluorophores using a single detection fiber in the confocal arm. A dichroic mirror separates the fluorescence from the two fluorophores, which is separately detected using a PMT array.

7.3 Conclusion

This chapter explored different designs for the modification of the excitation and detection system for the array confocal microscope presented in Chapter 4 to adapt that system for the detection of multiple fluorophores. One simple modification for the excitation system is to use multiple beam-splitting gratings optimized for a single wavelength to independently couple each excitation wavelength into multiple fibers. The multiple fibers are then combined and made collinear on the confocal arm for multi-wavelength excitation.

There are multiple choices to re-design the detection system for the detection of multiple fluorophores. Designs that utilize one or two detection fibers in the confocal arm are possible. Each has its own advantages and disadvantages.

The achromatic beam-splitting gratings described in Chapters 5 and 6 can also be used to modify the excitation system. While it is possible to use the achromatic grating

doublets described in Chapter 5, this doesn't yield a lot of advantages if coupling the light to single-mode fibers. However, using the scale-invariant achromatic gratings described in Chapter 6 yields multiple advantages. By using the SIA gratings, multiple wavelengths can be coupled into one single-mode fiber. This reduces the total number of single-mode fibers needed for coupling by half. Since both excitation wavelengths exit the same single-mode fiber, a dichroic mirror does not have to be used to combine the beams and make them collinear. Using a SIA grating would not only greatly simplify the design, but it would also save money, since fewer components are required.

8. CONCLUSION

8.1 Summary

Confocal microscopy is a powerful tool that allows for three-dimensional, high-resolution imaging inside scattering media. Chapter 1 describes the details of confocal microscopes and the current state of the technology. Since confocal microscopy is a beam-scanning method, the image acquisition speed is limited. By using an array of beams to scan the object of interest, the acquisition speed over that FOV is decreased by a factor of the number of beams used in scanning.

Chapter 2 explores a design for an array confocal microscope that overcomes the trade-off between NA and FOV. By utilizing an array of miniature objectives, the total field of view is increased without sacrificing NA. In this design, each objective works on-axis and is translated laterally to scan a sub-FOV. All miniature objectives simultaneously scan their respective sub-FOVs. The images from each objective are stitched together in post-processing to stitch all sub-FOV images into one large FOV image. While this design overcomes the trade-off between NA and FOV, one limitation is speed. Since the objective array has to be mechanically scanned, the acquisition speed is limited. Furthermore, the design is very sensitive to fabrication errors for diffraction-limited performance. For these reasons, the design was not fabricated.

A different design for an array confocal microscope is proposed and tested in Chapter 4 for whole brain imaging at speeds faster than current methods. Microscope objectives are not suited for whole brain imaging due to their small FOVs and short working distances. These limitations are overcome by designing a large objective with a 0.5 NA, 3 mm × 3 mm FOV, and 15.5 working distance. This objective provides sub-cellular

resolution over a much larger FOV than current microscopes, and the large working distance allows the bottom of a brain to be imaged without hitting the objective. To further speed up the image acquisition, a 3×3 array of beams is scanned in the sample plane using a galvo mirror to reduce the acquisition time over this FOV by a factor of 9. The brains to be imaged using this microscope are made optically transparent using a brain clarification method called CLARITY. This array confocal microscope is built and tested. The built microscope has an estimated resolution of 657 nm, and a working distance of 10 mm is demonstrated. A cleared brain is imaged using this microscope to demonstrate its potential for whole brain imaging.

Beam-splitting gratings are an extremely efficient method to split one beam into a number of equal efficiency beams. Beam-splitting gratings are instrumental in the design of the array confocal microscopes described in Chapters 2 and 4. Chapter 3 details a design process to optimize single wavelength beam-splitting diffraction gratings with reduced sensitivity to fabrication errors. Performance degradation due to errors in the fabricated height of the gratings and errors caused by the finite beam width of the maskless lithography tool is minimized using the proposed design method.

One disadvantage of the beam-splitting gratings in Chapters 2-4 is it is designed only for a single wavelength. The performance may be degraded significantly if the wavelength is not the designed wavelength. Therefore, achromatic gratings that utilize two grating layers are explored in Chapters 5 and 6. These achromatic grating doublets utilize two grating layer with different indices of refraction to provide the designed performance at exactly two wavelengths, which is analogous to how an achromatic lens is designed. Chapter 5 analyzes the performance an achromatic grating doublet to

fabrication errors in the lateral alignment of the grating layers and errors in the fabricated heights of the grating layers. An analysis of the lateral shift of the grating layers proved that grating profiles with constant first spatial derivatives are significantly less sensitive than continuous phase profiles. One possible problem with fabrication of these designs is the relatively large heights of the profiles. Therefore, a design method to reduce the heights of the grating layers for binary grating profiles is proposed that adds an integer multiple of 2π to the phase of the grating. While this method significantly reduces the height of the grating layers, the performance of this reduced height grating is degraded for the wavelengths between the designed wavelengths.

The achromatic grating doublets described in Chapter 5 assume the phase profile is identical at two wavelengths. While the performance of the grating is identical at two wavelengths, the diffraction angle at these two wavelengths is different. Non-equal diffraction angles after the gratings severely complicates coupling of the diffracted light into fibers. Chapter 6 doesn't make the assumption that the phase profiles are identical to demonstrate that gratings can be designed with designed performance at two wavelengths and equal diffraction angles at these two wavelengths. It is shown that the phase profiles at these two designed wavelengths do not have to be identical or even similar. The method described in Chapter 5 to reduce the height of the grating layers is explored, however, due to the structure of the grating layers, while not impossible, it is unlikely that the heights are reduced significantly by adding an integer multiple of 2π to the designed phase for a binary profile.

Finally, modifications to the array confocal microscope described in Chapter 4 are proposed to adapt the system for multiple wavelength excitation and detection in

Chapter 7. Some of these modifications take advantage of the achromatic beam-splitting gratings described in Chapters 5 and 6. These modifications allow for the increased speed provided by an array confocal microscope over a large volume. With multiple excitation wavelengths, multiple fluorophores in a sample can be studied simultaneously to probe even more information about the object of interest.

8.2 Future Directions

This dissertation has focused on array confocal microscopy, however, the concept of array microscopy is not limited to a confocal microscope. One particular imaging modality that may benefit from parallelization is two-photon microscopy. Two-photon microscopy is similar to confocal microscopy in that it provides optical sectioning inside scattering media, however, two-photon microscopy is capable of deeper penetration and reduced phototoxicity, since the excitation wavelengths are typically near infrared. Like confocal microscopy, two-photon microscopy is a point-scanning method, however, a detection pinhole is not necessary. Two-photon microscopy typically utilizes a high power near infrared laser, so a fluorophore is excited when two infrared photons are absorbed. A fluorophore is excited in two-photon microscopy when two infrared photons are almost simultaneously absorbed to provide the equivalent energy of a single photon with half the wavelength to the fluorophore. One problem is the probability of two photons being absorbed by a fluorophore almost simultaneously is extremely low. If a high power light source, such as a pulsed laser, is used, then the photon flux at the focus is high, which results in two-photon absorption. Since two-photon microscopy utilizes long wavelengths for excitation, the penetration depth is significantly deeper than confocal microscopy. Furthermore, since only the volume in which two photons are

absorbed is excited, the risk of photobleaching is significantly lower than in confocal microscopy.

One complication in designing an array two-photon microscope is designing an efficient beam-splitting method. Since pulsed lasers are typically required, the spectral bandwidth of the laser is larger than the lasers used in confocal microscopes. If a beam-splitting diffraction grating is used, the larger spectral bandwidth may cause not only lateral chromatic aberration for non-zero diffraction orders, but also the efficiency of the grating over the spectral bandwidth may not be constant. Another concern is the required power for the laser for an array two-photon microscope. Two-photon microscopes typically require high power lasers to ensure two photons are absorbed by a fluorophore, however, if the incident laser is split into N beams, then the overall power of the laser also decreases by a factor of N, which may not be high enough to ensure two photons are absorbed by the fluorophore. While there are challenges associated with an array two-photon microscope, a two-photon microscope that is designed similarly to the array confocal microscope described in Chapter 4 would have many benefits for any application that requires a large volume to be imaged inside scattering media.

Another possible direction for future work is the fabrication of the achromatic gratings described in Chapters 5 and 6. Due to the large heights associated with the achromatic gratings, fabrication is challenging, but not impossible. Further work can explore different design methods to achromatize beam-splitting gratings. The design methods should pay special attention to reducing the heights to be significantly smaller than the designs in Chapters 5 and 6.

9. REFERENCES

1. S.W. Hell and J. Wichmann, “Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy,” Optics Letters 19(11), 780-782 (1994).

2. T.A. Klar and S.W. Hell, “Subdiffraction resolution in far-field fluorescence microscopy,” Optics Letters 24(14), 954-956 (1999).

3. G. Moneron and S.W. Hell, “Two-photon excitation STED microscopy,” Optics Express 17(17), 14567-14573 (2009).

4. K.I. Willig, et al., “Nanoscale resolution in GFP-based microscopy,” Nature Methods 3(9), 721 (2006).

5. M. J. Rust, M. Bates, and X. Zhuang, “Sub-diffraction-limit imaging by stochastic optical
reconstruction microscopy (STORM),” Nature methods **3**(10), 793-796 (2006).

6. S. van de Linde, et al., “Direct stochastic optical reconstruction microscopy with
standard fluorescent probes,” Nature Protocols, **6**(7), 991-1009 (2011).

7. P.J. Keller, et al., “Reconstruction of zebrafish early embryonic development by scanned
light sheet microscopy,” Science **322**(5904), 1065-1069 (2008).

8. M. Heilemann, et al., “Subdiffraction-resolution fluorescence imaging with conventional
fluorescent probes,” Angewandte Chemie International Edition **47**(33), 6172-6176 (2008).

9. S.T. Hess, T.P. Girirajan, and M.D. Mason, “Ultra-high resolution imaging by
fluorescence photoactivation localization microscopy,” Biophysical Journal **91**(11), 4258-4272 (2006).

10. T.J. Gould, V.V. Verkhusha, and S.T. Hess, “Imaging biological structures with
fluorescence photoactivation localization microscopy,” Nature Protocols **4**(3), 291-308 (2009).

11. M. G. Gustafsson, “Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy,” Journal of Microscopy 198(2), 82-87 (2000).

12. L. Schermelleh, et al., “Subdiffraction multicolor imaging of the nuclear periphery with 3D structured illumination microscopy,” Science 320(5881), 1332-1336 (2008).

13. M. G. Gustafsson, et al., “Three-dimensional resolution doubling in wide-field fluorescence microscopy by structured illumination,” Biophysical Journal 94(12), 4957-4970 (2008).

14. L. Shao et al., “Super-resolution 3D microscopy of live whole cells using structured illumination,” Nature Methods 8(12), 1044-1046 (2011).

15. M. G. Gustafsson, “Nonlinear structured-illumination microscopy: wide-field fluorescence imaging with theoretically unlimited resolution,” Proceedings of the National Academy of Sciences of the United States of America 102(37), 13081-13086 (2005).

16. E. H. Rego, et al., “Nonlinear structured-illumination microscopy with a photoswitchable protein reveals cellular structures at 50-nm resolution,” Proceedings of the National Academy of Sciences 109(3), E135-E143 (2012).

17. M. Minksy, “Microscopy apparatus,” U.S. Patent No. 2,013,467(1961).

18. R. W. Cole, T. Jinadasa, and C. M. Brown, “Measuring and interpreting point spread functions to determine confocal microscope resolution and ensure quality control,” Nature Protocols 6 (12), 1929-1941 (2011).

19. S. Wilhelm, B. Gröbler, M. Gulch, and H. Heinz, *Confocal Laser Scanning Microscopy: Principles*, Carl Zeiss, Inc. Jena, Germany (2003).

20. V. Drazic, "Dependence of two- and three-dimensional optical transfer functions on pinhole radius in a coherent confocal microscope," J. Opt. Soc. Am. A 9, 725-731 (1992).

21. T. Wilson and A. R. Carlini, “Size of detector in confocal imaging system,” Opt. Lett. 12, 227-229 (1987).

22. C. J. R. Sheppard and S. Rehman, “Confocal Microscopy,” in Biomedical Optical Imaging Technologies Design and Applications, R. Liang, ed. (Springer, 2013).

23. H. Goldman, “Spaltlampenphotographie und –photometrie,” Ophthalmologica 98, 257-270 (1940).

24. M. Petran, M. Hadravsky, M. D. Egger, and Galambos, “Tandem-scanning reflected-light microscope,” *J. Opt. Soc. Am.* 58, 661-664 (1968).

25. Takeo Tanaami, Shinya Otsuki, Nobuhiro Tomosada, Yasuhito Kosugi, Mizuho Shimizu, and Hideyuki Ishida, "High-speed 1-frame/ms scanning confocal microscope with a microlens and Nipkow disks," Appl. Opt. 41, 4704-4708 (2002).

26. E. L. Botvinick, F. Li, S. Cha, D. A. Gough, Y. Fainman, and J. H. Price, “In vivo confocal microscopy based on the Texas Instruments digital micromirror device,” *Proc. SPIE* 3921, 12-19 (2000).

27. M. Kanai, “Condensing optical system, confocal optical system, and scaning confocal endoscope,” U.S. Patent Application No. 20050052753 (2005).

28. D. L. Dickensheets and G. S. Kino, “Micromachined scanning confocal optical microscope,” *Opt. Lett.* 21, 764-766 (1996).

29. E. J. Seibel and Q. Y. L. Smithwick, “Unique features of optical scanning single fiber endoscopy,” *Lasers Surg. Med.* 30(3), 177-183 (2002).

30. D. L. Dickensheets and G. S. Kino, “Silicon-micromachined scanning confocal optical microscope,” *J. Microelectromech. Syst.* 7, 38-47 (1998).

31. H. Ra, et al., "Three-dimensional in vivo imaging by a handheld dual-axes confocal microscope," Opt. Express 16, 7224-7232 (2008).

32. K. C. Maitland, H. J. Shin, H. Ra, D. Lee, O. Solgaard, and R. Richards-Kortum, “Single fiber confocal microscope with a two-axis gimbaled MEMS scanner for cellular imaging,” *Opt. Express* 14, 8604-8612 (2006).

33. H. Miyajima, et. al, “A MEMS electromagnetic optical scanner for a commercial confocal laser scanning microscope,” *J. Microelectromech. Syst.* 12, 243-251 (2003).

34. E. J. Seibel, Q. Y. J. Smithwick, C. M. Brown, and P. G. Reinhall, “Single fiber flexible endoscope: general design for small size, high resolution and wide field of view,” *Proc. SPIE* 4158, 29-39 (2001).

35. G. E. Rosman, B. C. Jones, R. A. Pattie, and C. G. Byrne, “Optical fiber scanning apparatus,” U.S. Patent 7 920 312, Apr. 5, 2011.

36. Z. Li and L. Fu, “Note: A resonant fiber-optic piezoelectric scanner achieves a raster pattern by combining two distinct resonances,” Rev. of Sci. Instrum. 83, 086102 (2012).

37. H. Mansoor, H. Zeng, I. T. Tai, J. Zhao, and M. Chiao, “A Handheld Electromagnetically Actuated Fiber Optic Raster Scanner for Reflectance Confocal Imaging of Biological Tissues,” IEEE Transactions on Biomedical Engineering 60, 1431-1438.

38. A. F. Gmitro and D. Aziz, “Confocal microscopy through a fiber-optic imaging bundle,” *Opt. Lett.* 18, 565-567 (1993).

39. C. P. Lin and R. H. Webb, “Fiber-coupled multiplexed confocal microscope,” *Opt. Lett.* 25, 954-956 (2000).

40. P. M. Lane, A. L. P. Dlugan, R. Richards-Kortum, and C. E. MacAulay, “Fiber-optic confocal microscopy using a spatial light modulator,” *Opt. Lett.* 25, 1780-1782 (2000).

41. M. B. Sinclair, D. M. Haaland, J. A. Timlin, and H. D. T. Jones, "Hyperspectral confocal microscope," Appl. Opt.45, 6283-6291 (2006).

42. D. M. Haaland, et. al, “Hyperspectal confocal fluorescence imaging: Exploring alternative multivariate curve resolution approaches,” Appl. Spectrosc. 63(3), 271-279 (2009).

43. G. Lu and B. Fei, “Medical hyperspectral imaging: a review,” J. Biomed. Opt. 19(1), 010901 (2014).

44. Olympus, “PLN objectives,” http://www.olympus-lifescience.com/en/objectives/pln/

45. A. Orth and K. Crozier, "Gigapixel fluorescence microscopy with a water immersion microlens array," Opt. Express 21, 2361-2368 (2013).

46. H. J. Tiziani, R. Achi, R. N. Krämer, and L. Wiegers, "Theoretical analysis of confocal microscopy with microlenses," Appl. Opt.35, 120-125 (1996).

47. H. J. Tiziani and H. Uhde, "Three-dimensional analysis by a microlens-array confocal arrangement," Appl. Opt. 33, 567-572 (1994).

48. A. Orth and K. Crozier, "Microscopy with microlens arrays: high throughput, high resolution and light-field imaging," Opt. Express 20, 13522-13531 (2012).

49. M. Hillenbrand, et al, "Chromatic confocal matrix sensor with actuated pinhole arrays," Appl. Opt. 54, 4927-4936 (2015).

50. M. Ulrich, and S. Lange-Asschenfeldt,“In vivo confocal microscopy in dermatology from research to clinical application.” Journal of Biomedical Optics 18(6), 061212 (2013).

51. A. L. Clark, et al., “Confocal microscopy for real-time detection of oral cavity neoplasia,” Clin Cancer Res 9, 4714-4721 (2013).

52. Physik Instrumente, “Compact linear positioning stage,”
http://www.pi-usa.us/products/PDF_Data/M663_Precision_Positioning_Stage.pdf

53. S. Pacheco, J. F. Brand, M. Zaverton, T. Milster, and R. Liang, "Sensitivity analysis and optimization method for the fabrication of one-dimensional beam-splitting phase gratings," Opt. Express **23**, 11771-11782 (2015).

54. M. Gruber, M. Jarczynski, R. Kerssenfischer, and J. Jahns, "Fan-out techniques using planar-integrated free-space optics," in Frontiers in Optics, OSA Technical Digest (CD) (Optical Society of America, 2003), paper WI4.

55. J. Turunen, A. Vasara, J. Westrholm, G. Jin, and A. Salin, "Optimization and fabrication of beam splitters," J. Phys. D **21**, s202-s205 (1988).

56. M. Dames, R. Dowling, P. McKee, and D. Wood, "Efficient optical elements to generate intensity weighted spot arrays: design and fabrication," Appl. Opt. **30**, 2685-2691 (1991).

57. H. Dammann, and K. Görtler, "High-efficiency in-line multiple imaging by means of multiple phase holograms," Optics Communications **3**, 312-315 (1971).

58. L. Boivin, "Multiple imaging using various types of simple phase gratings," Appl. Opt. **11**, 1782-1792 (1972).

59. Yu. V. Miklyaev, A. Krasnaberski, M. Ivanenko, A. Mikhailov, W. Imgrunt, L. Aschke, and V. N. Lissotschenko, "Efficient diffractive optical elements from glass with continuous surface profiles," Proc. SPIE 7913, 79130B (2011).

60. A. Krasnaberski, Y. Miklyaev, D. Pikhulya, L. Kleinschmidt, W. Imgrunt, M. Ivanenko, and V. Lissotschenko, "Efficient beam splitting with continuous relief DOEs and microlens arrays," Proc. SPIE 8236, 823609 (2012).

61. M. Jarczynski, et al., "Monolithic optical components for splitting of high-power beams," Proc. SPIE 8490, 84900Q (2012).

62. L. Romero and F. Dickey, "The mathematical theory of laser beam-splitting gratings," Prog. Opt. **54**, 319-386 (2010).

63. L. Romero and F. Dickey, "Theory of optimal beam splitting by phase gratings. I. One-dimensional gratings," J. Opt. Soc. Am. A **24**, 2280-2295 (2007).

64. L. Romero and F. Dickey, "Theory of optimal beam splitting by phase gratings. II. Square and hexagonal gratings," J. Opt. Soc. Am. A **24**, 2296-2312 (2007).

65. Y. Miklyaev, W. Imgrunt, V. Pavelyev, D. Kachalov, T. Bizjak, L. Aschke, V. Lissotschenko, "Novel continuously shaped diffractive optical elements enable high efficiency beam shaping," Proc. SPIE 7640, 764024 (2010).

66. T. Lizotte, R. Rosenberg, and O. Ohar, "Actual performance vs. modeled performance of diffractive beam splitters," Proc. SPIE 5876, 58760I (2005).

67. M. Zaverton, J. Sierchio, Y. Kim, D. Hansen, W. Bletcher, J. Tamkin, and T. Milster, "Maskless lithography tool for fabrication and inspection of diffractive optical elements and computer generated holograms," in International Optical Design Conference and Optical Fabrication and Testing, OSA Technical Digest (CD) (Optical Society of America, 2010), paper OTuC5.

68. J. W. Goodman, *Introduction to Fourier optics* (McGraw Hill, 1968).

69. MATLAB R2013a, MathWorks, Inc., Natick, Massachusetts, United States.

70. K. D. Micheva, and S. J. Smith, "Array tomography: a new tool for imaging the molecular architecture and ultrastructure of neural circuits," Neuron 55, 25–36 (2007).

71. D. D. Bock, et al., "Network anatomy and *in vivo* physiology of visual cortical neurons," Nature 471, 177–182 (2011).

72. T. Ragan, et al., "Serial two-photon tomography for automated *ex vivo* mouse brain imaging," Nature Methods 9, 255–258 (2012).

73. J. DeFelipe, “From the connectome to the synaptome: an epic love story,” Science 330,1198–1201 (2010).

74. K. Chung, et al, “Structural and molecular interrogation of intact biological systems,” Nature 497, 333-337 (2013).

75. E. A. Susaki, et al, “Whole-brain imaging with single-cell resolution using chemical cocktails and computational analysis,” Cell 157, 726-739 (2014).

76. H. U. Dodt, et al.,“Ultramicroscopy: three-dimensional visualization of neuronal networks in the whole mouse brain,” Nat. Methods 4, 331–336 (2007).

77. H. Hama, et al., “Scale: a chemical approach for fluorescence imaging and reconstruction of transparent mouse brain,” Nat. Neurosci. *14*, 1481–1488 (2011).

78. V. Rankov, et al., “An algorithm for image stitching and blending,” Proc. SPIE 5701, Three-Dimensional and Multidimensional Microscopy: Image Acquisition and Processing XII, 190, (2005).

79. G. Cox, and C. J. R. Sheppard, “Practical limits of resolution in confocal and non-linear microscopy,” Microsc. Res. Tech. 83, 18-22 (2004).

80. S. Pacheco, T. Milster, and R. Liang, “Analysis of grating doublets for achromatic beam-splitting,” Opt. Express **23**, 22939-22952(2015).

81. Y. Arieli, S. Noach, S. Ozeri, and N. Eisenberg, “Design of diffractive optical elements for multiple wavelengths,” Appl. Opt. **37**, 6174-6177 (1998).

82. S. Noach, N. Eisenberg, and Y. S. Arieli, “Achromatic diffractive optical element,” Proc. SPIE 3778, Gradient Index, Miniature, and Diffractive Optical Systems, 151 (1999).

83. L. Yang, O. Cui, T. Liu, and C. Xue, "Effects of manufacturing errors on diffraction efficiency for multilayer diffractive optical elements," Appl. Opt. **50**, 6128-6133 (2011).

84. L. Gao, S. To, H. Yang, X. Nie, T. Liu, and C. Xue, "Effect of assembling errors on the diffraction efficiency for multilayer diffractive optical elements," Appl. Opt. **53**, 7341-7347 (2014).

85. T. R. M. Sales and D. H. Raguin, “Multiwavelength operation with thin diffraction elements,” Appl. Opt. **38**, 3012-3018 (1999).

86. G. I. Greisukh, E. A. Bezus, D. A. Bykov, E. G. Ezhov, and S. A. Stepanov, “Suppression of the spectral selectivity of two-layer phase-relief diffraction structures,” Optics and Spectroscopy **106**(4), 621-626 (2009).

87. Y. Ogura, N. Shirai, J. Tanida, and Y. Ichioka, "Wavelength-multiplexing diffractive phase elements: design, fabrication, and performance evaluation," J. Opt. Soc. Am. A **18**, 1082-1092 (2001).

88. M. W. Farn and J. W. Goodman, "Diffractive doublets corrected at two wavelengths," J. Opt. Soc. Am. A **8**, 860-867 (1991).