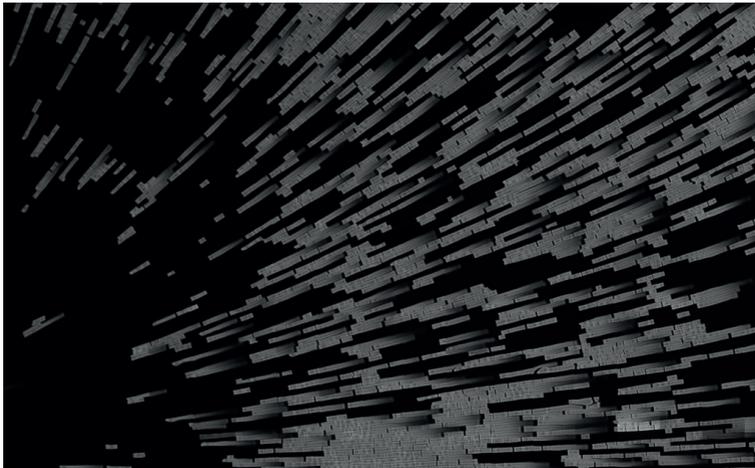

Zheng Li

Application of diffractive lens arrays in confocal microscopy



Schriftenreihe Automatische Sichtprüfung und Bildverarbeitung | Band 19

Zheng Li

**Application of diffractive lens arrays
in confocal microscopy**

Schriftenreihe Automatische Sichtprüfung und Bildverarbeitung
Band 19

Herausgeber: Prof. Dr.-Ing. habil. Jürgen Beyerer

Lehrstuhl für Interaktive Echtzeitsysteme
am Karlsruher Institut für Technologie

Fraunhofer-Institut für Optronik, Systemtechnik
und Bildauswertung IOSB

Application of diffractive lens arrays in confocal microscopy

by
Zheng Li

Karlsruher Institut für Technologie
Lehrstuhl für Interaktive Echtzeitsysteme

Application of diffractive lens arrays in confocal microscopy

Zur Erlangung des akademischen Grades eines Doktor-Ingenieurs
von der KIT-Fakultät für Informatik des Karlsruher Instituts für
Technologie (KIT) genehmigte Dissertation

von Zheng Li, M.Sc.

Tag der mündlichen Prüfung: 26. November 2021
Erster Gutachter: Prof. Dr.-Ing. habil. Jürgen Beyerer
Zweiter Gutachter: Prof. Dr. rer. nat. Uli Lemmer

Impressum



Karlsruher Institut für Technologie (KIT)
KIT Scientific Publishing
Straße am Forum 2
D-76131 Karlsruhe

KIT Scientific Publishing is a registered trademark
of Karlsruhe Institute of Technology.
Reprint using the book cover is not allowed.

www.ksp.kit.edu



*This document – excluding parts marked otherwise, the cover, pictures and graphs –
is licensed under a Creative Commons Attribution-Share Alike 4.0 International License
(CC BY-SA 4.0): <https://creativecommons.org/licenses/by-sa/4.0/deed.en>*



*The cover page is licensed under a Creative Commons
Attribution-No Derivatives 4.0 International License (CC BY-ND 4.0):
<https://creativecommons.org/licenses/by-nd/4.0/deed.en>*

Print on Demand 2022 – Gedruckt auf FSC-zertifiziertem Papier

ISSN 1866-5934
ISBN 978-3-7315-1188-5
DOI 10.5445/KSP/1000145608

Abstract

Confocal microscopy has been one of the most commonly used methods for surface measurement, and it has also been the gold standard in life science for many years. In traditional confocal microscopes, high resolution and large fields of view (FOVs) are difficult to achieve simultaneously by the microscope objectives, which limit the space-bandwidth products of the optical systems. To solve the problem, diffractive lens arrays (DLAs) with overlapping apertures have been proposed in previous research. They can produce spots with high numerical apertures (NAs) in a dense grid over a large area. Combined with low-NA objectives, the tiny spots can be used to scan the whole sample with a much higher resolution than the objectives. However, they have only been demonstrated in transmitted-light microscopes, which significantly limits their applications.

In this work, two new DLA concepts are proposed, which can be used in reflected-light microscopes. The first one is the See-through DLA. It is designed by the superposition of a plane wave component to increase the zero-order diffraction efficiency, which makes it more transparent with less disturbance in the imaging path. The second one is the Direct-imaging DLA. It is designed by the superposition of a second lens component and it acts exactly as an array of finite-conjugate objectives. A physically based simulation framework has been established for the design and optimization of the DLAs. Different diffraction propagation methods are compared and their sampling conditions are discussed.

Based on the proposed concepts, prototypes are manufactured to validate the performance of the DLAs. The See-through DLAs have been demonstrated to generate spots with 0.78 NA at a wavelength of 785 nm and 0.83 NA at a wavelength of 488 nm. Laterally, a spatial frequency up to $2048 \text{ lp}\cdot\text{mm}^{-1}$ has

been measured by the DLAs at 488 nm. Axially, an average height of 960.6 nm with a standard deviation of 49.2 nm has been measured for a calibrated step height target with a nominal height of 925.5 nm. By interference with the reflected conjugate waves, an average height of 904.7 nm with a standard deviation of 9.7 nm has been shown in the experiment. Fluorescence imaging has also been successfully demonstrated with the DLAs by measurement of the fluorescent beads, and an imaging NA of 0.83 has been achieved within an FOV of 3 mm \times 3 mm, which results in a space-bandwidth product of around 300 megapixels.

The Direct-imaging DLAs have been demonstrated to generate spots with an NA of 0.68 at a wavelength of 785 nm. Laterally, a spatial frequency of 1448 lp \cdot mm⁻¹ has been measured at 785 nm. Axially, by measurement of the step height target, an average height of 917.5 nm with a standard deviation of 49.9 nm has been achieved.

Both kinds of DLAs have shown capabilities of high-resolution measurement over a large area. They have overcome the limitations in the previous research and realized multi-spot confocal imaging in surface metrology and fluorescence microscopy.

Kurzfassung

Die konfokale Mikroskopie ist eine der am häufigsten verwendeten Methoden zur Oberflächenmessung. Sie ist auch seit vielen Jahren das Standardverfahren in den Biowissenschaften. In traditionellen konfokalen Mikroskopen sind hohe Auflösung und große Sichtfelder durch die Mikroskopobjektive nur schwer gleichzeitig zu realisieren, was die Orts-Bandbreite-Produkte der optischen Systeme begrenzen. Um das Problem zu lösen, wurden in der bisherigen Forschung die diffraktiven Linsenarrays (DLAs) mit überlappenden Aperturen vorgeschlagen. Sie können Spots mit hoher numerischer Apertur (NA) in einem dichten Gitter über einen großen Bereich erzeugen. Durch die Verwendung von Objektiven mit niedrigen NAs können die winzigen Spots genutzt werden, um die gesamte Probe mit einer viel höheren Auflösung zu scannen, verglichen mit der kleinen NA der Objektive. Bislang wurden sie jedoch nur in Transmissionsmikroskopen demonstriert, was ihre Anwendungsmöglichkeiten deutlich einschränkt.

In dieser Arbeit werden zwei neue DLA-Konzepte vorgeschlagen, die in Reflexionsmikroskopen eingesetzt werden können. Das erste ist die See-through-DLA. Sie wird durch Überlagerung einer ebenen Wellenkomponente entworfen, um die Effizienz in der nullten Beugungsordnung zu erhöhen, was sie transparenter macht, mit weniger Störungen im Abbildungspfad. Die Zweite ist die Direct-imaging DLA. Sie wird durch Überlagerung einer zweiten Abbildungskomponente konstruiert und verhält sich genau wie ein Array von endlich korrigierten Objektiven. In dieser Arbeit wurde eine physikalisch basierte Simulationsrahmen für das Design und die Optimierung der DLAs entworfen. Es werden verschiedene Beugungsausbreitungsmethoden verglichen und deren Abtastbedingungen diskutiert.

Basierend auf den vorgeschlagenen Konzepten wurden Prototypen hergestellt, um die Leistung der DLAs zu testen. Es wurde gezeigt, dass die See-through-DLAs Spots mit 0.78 NA bei einer Wellenlänge von 785 nm und 0.83 NA bei einer Wellenlänge von 488 nm erzeugen. Zusätzlich wurde von den DLAs eine Ortsfrequenz bis zu $2048 \text{ lp}\cdot\text{mm}^{-1}$ bei 488 nm gemessen. Axial wurde eine mittlere Höhe von 960.6 nm mit einer Standardabweichung von 49.2 nm für eine kalibrierte Stufenhöhenprobe mit einer Höhe von 925.5 nm gemessen. Durch Auswertung der Interferenz mit den reflektierten konjugierten Wellen wurde im Experiment eine mittlere Höhe von 904.7 nm mit einer Standardabweichung von 9.7 nm gemessen. Die Fluoreszenzabbildung wurde ebenfalls erfolgreich mit den DLAs durch Messung der fluoreszierenden Kügelchen demonstriert, und es wurde eine Abbildungs-NA von 0.83 bei einem Sichtfeld von $3 \text{ mm} \times 3 \text{ mm}$ erreicht, was zu einem Orts-Bandbreite-Produkt von etwa 300 Megapixeln führt.

Es wurde gezeigt, dass die Direct-imaging DLAs Spots mit einer NA von 0.68 bei einer Wellenlänge von 785 nm erzeugen. Zusätzlich wurde eine Ortsfrequenz von $1448 \text{ lp}\cdot\text{mm}^{-1}$ bei 785 nm gemessen. Axial wurde bei der Messung der Stufenhöhenprobe eine mittlere Höhe von 917.5 nm mit einer Standardabweichung von 49.9 nm erreicht.

Beide DLAs haben gezeigt, dass sie hochauflösende Messungen über eine sehr großes Sichtfeld durchführen können. Und sie haben die Einschränkungen in dem bisherigen Stand der Forschung überwunden und die konfokale Multispot-Bildgebung für die Oberflächenmessung und Fluoreszenzmikroskopie zugänglich gemacht.

Acknowledgements

I would like to send my deepest gratitude to Prof. Dr.-Ing. habil. Jürgen Beyerer, who has offered me the opportunity to work at the Vision and Fusion Laboratory (IES) of Karlsruhe Institute of Technology (KIT) in cooperation with Fraunhofer Institute of Optronics, System Technologies and Image Exploitation (IOSB). Without his persistent guidance and support, this work would not have been possible. I would also like to thank Prof. Dr. rer. nat. Uli Lemmer for being the second reviewer to give his valuable comments and suggestions on this dissertation.

I would like to express my special thanks to Dr.-Ing. Miro Taphanel, for his advice and help during my career. His creativity and enthusiasm have inspired me in many ways. Furthermore, I would like to thank all the members at IES, including Dr. Alexey Pak, Dr.-Ing. Chengchao Qu, Dr.-Ing. Johannes Meyer, Dr. Tim Zander, Ankush Meshram, Julius Krause, Patrick Philipp, Chengzhi Wu, and especially Dr.-Ing Ding Luo and Chia-Wei Chen, for the valuable talks, insightful discussions and heartfelt support, both as friends and colleagues. I am also grateful to all my colleagues at the Visual Inspection department (SPR) in IOSB, in particular Prof. Dr.-Ing. Thomas Längle, Dr.-Ing. Matthias Hartrumpf, Christian Negara, Christian Kludt, Benjamin-Alexander Timmerbeil, Dr. Wolfgang Melchert, Kai Niedernberg, and Alexander Enderle. It is a real pleasure to work with them.

I owe a debt of appreciation to my fiancée, Chenxi Jin, for her love and devotion during sunny and rainy days. At last, I would like to express the sincerest thanks to my parents, whose unconditional dedication have always helped me to overcome the difficulties in my life.

Karlsruhe, July 2021

Zheng Li

Contents

Abstract	i
Kurzfassung	iii
Acknowledgements	v
Notation	ix
Acronyms	xv
1 Introduction	1
1.1 Motivation	1
1.2 Related work	3
1.3 Main contributions	7
1.4 Thesis outline	9
2 Theoretical Background	11
2.1 Diffraction simulation	11
2.1.1 Maxwell's equations	12
2.1.2 Scalar approximation and solutions	16
2.1.3 Numerical implementation	27
2.2 Theory of confocal microscopy	38
2.2.1 Focusing by a thin lens	38
2.2.2 PSF of a thin lens	41
2.2.3 PSF of a confocal scanning microscope	45
3 Design and Simulation	53
3.1 See-through DLA design	53

3.1.1	System overview	54
3.1.2	Design process	57
3.1.3	Simulation and optimization	61
3.2	Direct-imaging DLA design	66
3.2.1	System overview	67
3.2.2	Design process	68
3.2.3	Simulation and optimization	71
4	Experiment Results	79
4.1	Spot characterization	80
4.1.1	Spot measurement of See-through DLAs	81
4.1.2	Spot measurement of Direct-imaging DLAs	88
4.2	Lateral measurement	91
4.2.1	Lateral measurement by See-through DLAs	91
4.2.2	Lateral measurement by Direct-imaging DLAs	97
4.3	Axial measurement	100
4.3.1	Axial response of See-through DLAs	100
4.3.2	Axial response of Direct-imaging DLAs	102
4.3.3	Interference measurement by See-through DLAs	104
4.3.4	Measurement of a step height target	106
4.4	Fluorescence measurement	110
5	Conclusions	119
5.1	Summary of the work	119
5.2	Outlook	121
	Bibliography	125
	Publications	139
 Appendix		
A	Kirchhoff's Diffraction Formulation	143

Notation

Conventions

Scalars	italic Latin letters	<i>u, U</i>
	italic Greek letters	<i>α, β</i>
Integers	upright Latin letters	m, M
Vectors & Matrices	bold Latin letters	h, H
Operators	calligraphic Latin uppercase letters	<i>\mathcal{F}, \mathcal{S}</i>

Operators

*	convolution of two signals
\circledast	circular convolution of two discrete signals
\odot	element-wise multiplication of two matrices with identical dimensions
∇f	gradient of the function f
$\nabla \cdot \mathbf{F}$	divergence of the vector field \mathbf{F}
$\nabla \times \mathbf{F}$	curl of the vector field \mathbf{F}
$\nabla^2 \mathbf{F}$	Laplace operator, applied to a vector field \mathbf{F}
$\angle(\mathbf{a}, \mathbf{b})$	angle between two vectors \mathbf{a} and \mathbf{b}
$(\cdot)^*$	complex conjugate of a variable
$ \cdot $	absolute value of a variable
$\lfloor x \rfloor$	greatest integer $\leq x$

$\arg(\cdot)$	argument (phase angle) of a complex-valued quantity
$\arg \max$	argument of the maximum
$\text{DFT}\{\mathbf{h}\}$	discrete Fourier transform of a matrix \mathbf{h}
$\text{IDFT}\{\mathbf{h}\}$	inverse discrete Fourier transform of a matrix \mathbf{h}
$\text{FFT}\{\mathbf{h}\}$	fast Fourier transform of a matrix \mathbf{h}
$\text{IFFT}\{\mathbf{h}\}$	inverse fast Fourier transform of a matrix \mathbf{h}
$\mathcal{F}\{f(x,y)\}$	Fourier transform of a function $f(x,y)$
$\mathcal{F}^{-1}\{f(x,y)\}$	inverse Fourier transform of a function $f(x,y)$
$\text{mod}(x, N)$	modulo operation of the variable x to N
$S\{u, k, d\}$	diffraction propagation of the field u with a wave number k by the Rayleigh-Sommerfeld integral through a distance d .
$\mathbf{u}[m,n]$	the $[m,n]$ 'th element of the matrix \mathbf{u} .

Greek Symbols

α, β, γ	spatial frequency components
$\delta(x)$	Dirac delta function
δz	axial displacement from the focal plane
$\Delta d, \Delta x, \Delta y$	spatial or sampling intervals
Δf_x	frequency intervals
ϵ_0	vacuum permittivity
$\epsilon(\omega)$	relative permittivity or dielectric constant
θ	half angle of the light cone of a focused spot
λ	wavelength
μ_0	vacuum permeability
ξ, η	lateral coordinates on the lens
ρ	normalized radius
ρ_f	free charge density

$\sigma(\omega)$	electric conductivity
σ_S	standard deviation
$\Sigma, \bar{\Sigma}$	surfaces on the screen
ϕ_D, ϕ_S	binary phases of the phase masks
$\varphi_{ASM}, \varphi_{RSI}$	oscillating phase terms
$\chi(\omega)$	electric susceptibility
ω	angular frequency

Latin Symbols

a	radius of a circular aperture
a_0, a_1, a_2	line fitting parameters
B	binarization factor
$\mathbf{B}(\mathbf{r}, t)$	magnetic induction
c_0	speed of light in vacuum
C_1, C_2	arbitrary constants
d_0, d_1, d_2	distances
$d_S, d_{S, \text{mean}}$	altitude of the step height target
$\mathbf{D}(\mathbf{r}, t)$	electric displacement
$\mathbf{E}(\mathbf{r}, t)$	electric field
f	focal length
f_s	sampling frequency
f_{max}	maximum frequency of a signal
G, G_+, G_-	Green's functions
$h(x, y)$	amplitude point spread function
$h_z(x, y)$	impulse response function
\mathbf{h}_z	discretized impulse response function in matrix form
$\overset{\circ}{\mathbf{h}}_z$	extended impulse response function matrix

$\mathbf{H}(\mathbf{r}, t)$	magnetizing field
$H_z(\alpha, \beta)$	transfer function
$\mathring{\mathbf{H}}_z$	extended transfer function matrix
$I(x, y)$	intensity distribution
$\mathbf{j}(\mathbf{r}, t)$	conduction current density
$\mathbf{j}_f(\mathbf{r}, t)$	free current density
$J_0(x)$	zero-order Bessel function of the first kind
$J_1(x)$	first-order Bessel function of the first kind
\mathbf{k}	wave vector
$k(\omega)$	wave number
k_0	wave number in air
L	width of the step height target
m, n, M, N	integer indices
$\mathbf{M}(\mathbf{r}, t)$	magnetization field
M_a	magnification
n	refractive index
\mathbf{n}	unit vector in the normal direction of a surface
N_F	Fresnel number
$O(N)$	computational complexity of N
p, q	integer indices
$\mathbf{P}(\mathbf{r}, t)$	polarization field
$P(x, y)$	circular aperture function
r	radius or lateral distance to the origin
$\mathbf{r}, \mathbf{r}', \bar{\mathbf{r}}$	coordinates in vector form
R	radius to the boundary
s_1, s_2	arbitrary integers
$t(x, y)$	transmittance function
T	pitch of an array
$u(\mathbf{r}, \omega)$	scalar field

\mathbf{u}	discretized scalar field in matrix form
$\mathring{\mathbf{u}}_z$	extended scalar field matrix
$U(\mathbf{k}, \omega)$	scalar field in the spatial frequency domain
\mathbf{U}	discretized scalar field in the spatial frequency domain in matrix form
\mathbf{v}_ρ	volumetric charge velocity
v	radial optical coordinate
w	axial optical coordinate
W_D	ratio of the lens elements
W_S	ratio of the plane wave component
x, y, z	coordinates

Acronyms

3D	three-dimensional
AFWHM	axial full width at half maximum
ASM	angular spectrum method
CLSM	confocal laser scanning microscope
DFT	discrete Fourier transform
DLA	diffractive lens array
DOE	diffractive optical element
FDTD	finite-element time-domain
FFT	fast Fourier transform
FOV	field of view
FWHM	full width at half maximum
HWHM	half width at half maximum
KIT	Karlsruhe Institute of Technology
LFWHM	lateral full width at half maximum
MLA	micro lens array
MTF	modulation transfer function
NA	numerical aperture
OTF	optical transfer function
PBR	peak-to-background ratio
RSI	Rayleigh-Sommerfeld integral
SBP	space-bandwidth product
SNR	signal-to-noise ratio

1 Introduction

1.1 Motivation

Since its first invention by Minsky in the 1950s [Min61], confocal microscopy has become one of the most important advances of light microscopy in the past decades [Wil09]. Confocal microscopes have the advantages over conventional wide-field microscopes in resolution and contrast, making them widely used in biomedical imaging and industrial metrology.

The setup of a typical reflected-light scanning confocal microscope is shown in Figure 1.1 (a). A point light source, usually a laser, is collimated and focused to a spot by an objective. The spot illuminates a tiny part of the sample. Then, it is imaged by the objective onto the detector. A small pinhole is used to block

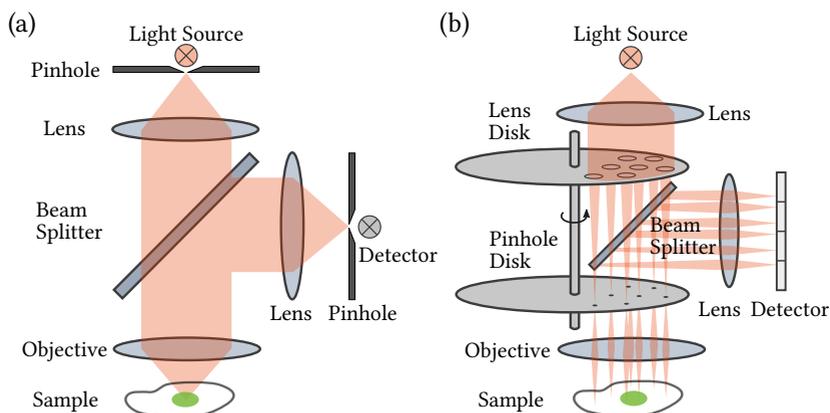


Figure 1.1: Schematics of common confocal microscope setups. (a) A single-spot confocal scanning microscope. (b) A spinning-disk confocal microscope.

the out-of-focus light scattered by the other parts of the sample. In this way, the sample is scanned by the spot and a sharper image can be reconstructed. Confocal microscopes can image the sample with higher resolution than conventional wide-field microscopes in both the lateral and the axial directions. The pinhole can be replaced by a pinhole array to increase the measurement speed by multi-spot scanning. For example, a typical Yokogawa spinning-disk confocal microscope is shown in Figure 1.1 (b) [Fav92, Ich96]. A Nipkow disk [Nip84] rotates along the axis and the sample is scanned by multiple spots which are produced by the pinholes on the disk. Comparing to the single-spot scanning, it can achieve a higher measurement speed.

The spatial resolution of a confocal microscope is determined by the numerical aperture (NA) of the objective, which is defined as

$$\text{NA} = n \sin \theta, \quad (1.1)$$

where n is the refractive index of the medium and θ is the half angle of the light cone focused by the objective. According to the well-known resolution limit, which was first described by Abbe in 1873 [Abb73] and later refined by Lord Rayleigh in 1879 [FRS79], the smallest resolvable distance between two points which are imaged by a microscope is inversely proportional to the NA of the objective

$$\Delta d = 0.61 \frac{\lambda}{\text{NA}}, \quad (1.2)$$

where Δd is the minimum resolvable distance and λ is the wavelength. Such a relation is also widely known as the Rayleigh criterion.

Subsequently, high-NA objectives are generally required in order to image the sample with high resolution, while they usually also have large magnifications, which lead to small fields of view (FOVs). The relation between the resolution and the FOV of an optical system can be described by the space-bandwidth product (SBP), which is, in the two-dimensional case, defined as

$$\text{SBP} = \frac{\text{FOV}}{(0.5\Delta d)^2}, \quad (1.3)$$

where FOV is the total imaging area [Loh96]. The SBP of an optical system has the unit in pixels and represents the total number of resolvable pixels by the optical system. It describes the information throughput of the optical system. Typically, for example, a $40\times$ 0.75 NA objective (Olympus UPlan FLN $40\times/0.75$) has a field diameter of 0.66 mm and a resolution of $0.41\ \mu\text{m}$, which leads to an SBP of around 8 megapixels. The commercial microscope objectives usually have the SBPs in megapixels [Bia17], independent of their magnifications or NAs. Although one can scale up the size of the lens and compensate for the aberrations by increasing the number of lens elements to increase the FOV and the SBP, the resulting objectives, e.g., lithography lenses, are extremely expensive and difficult to manufacture, which makes them impractical for microscopy applications [Zhe16].

However, with the development of nanotechnologies nowadays, there are increasing demands for precise measurement of small structures over large areas, such as semiconductor wafers and meta-surfaces. These applications would require, for example, sub-micron resolution over an area in centimeters, which presents challenges for traditional confocal microscopy due to the aforementioned trade-offs between resolution and FOVs. Therefore, it is desirable to develop new technologies to overcome such limitations and increase the SBPs in order to perform high-resolution large-area measurement with a fast speed and at a low cost.

1.2 Related work

Single-spot scanning confocal microscopes, which are often referred to as confocal laser scanning microscopes (CLSMs), are the most common setups in confocal microscopy. They have relatively simple structures, making it flexible to vary the pinhole sizes, to change the scanning trajectories and to integrate multiple detectors for simultaneous multi-wavelength detection. CLSMs usually use resonant galvanometer mirrors as scanners and photo-multiplier tubes or avalanche photodiodes as detectors. Although recent efforts have been made to use micro mirrors [Arr10] or polygonal mirrors [Cho13] to accelerate the scanning, the overall measurement speed is still limited

by the detector noise and sensitivity [Ore14] in certain applications such as fluorescence microscopy with low light intensities.

Conversely, multi-spot scanning confocal microscopes utilize parallel scanning with camera sensors which have higher quantum efficiencies to realize higher frame rates [Too06]. Traditionally, the projection of multiple spots is achieved by pinhole arrays or beam splitting in front of the objectives. The most common type is the spinning-disk setup which was first introduced by Petran et al. [Pet68] and later modified by other researchers [Fav92, Ich96, Ino02, Tan02]. Besides, a number of papers [Pac17, Ada18, Bes18] and patents [Lun08, Abr16, Sch20, Mat21] have used various kinds of diffractive optical elements (DOEs), e.g., Dammann gratings [Jah89], to split the illumination beam into multiple beams with different incident angles. Then, the beams are also scanned by the traditional galvanometer mirror scanners as the CLSMs. Nevertheless, such multi-spot confocal microscope configurations still rely on the objectives to project spots. Consequently, they remain to suffer from the trade-offs between FOVs and NAs, and thus they are limited in SBPs.

In previous research, a variety of array illuminators [Loh92] have shown the capabilities to focus the incident light into arrays of spots. The spot arrays can be used to scan the sample and perform confocal imaging like Figure 1.2 (a) shows. The projected spots illuminate the sample and are imaged onto a camera sensor. The pixels on the sensor can be treated as pinholes and confocal images are obtained by scanning of either the sample or the spot array. Among them, the array illuminators based on the Talbot effect, micro lens arrays (MLAs) and diffractive lens arrays (DLAs) can generate spots by themselves and do not need extra lenses for focusing. Their sizes and FOVs are scalable, and thus they have the potential to address the restriction between FOVs and resolution in confocal microscopy.

For example, Pang et al. have demonstrated [Pan12, Pan13, Sun16] their Talbot-effect-based multi-spot scanning microscopes which produce spots with NAs of 0.2, and they have measured a spatial resolution up to $417 \text{ lp}\cdot\text{mm}^{-1}$ with an FOV of $3.9 \text{ mm} \times 3.5 \text{ mm}$ at a wavelength of 488 nm. Orth et al. [Ort12, Ort14] have built fluorescence microscopes based on MLAs with measured NAs of around 0.35 and an FOV of $5.5 \text{ mm} \times 5.5 \text{ mm}$ at a wavelength of 532 nm.

However, in theory, the Talbot effect is derived from Fresnel's diffraction with paraxial approximation [Lat92], while MLAs suffer from aberrations due to the difficulties in controlling the surface shape. Therefore, it is hard for both of the technologies to realize high NAs. On the contrary, the micro structures of the DLAs are precisely made by lithography and they can produce diffraction-limited spots which are almost free of geometrical aberrations. For example, Zhu et al. have presented a supercritical DLA which produces spots with an equivalent NA of 0.83, a pitch of $200\ \mu\text{m}$ and a focal length of about $64\ \mu\text{m}$ at a wavelength of $633\ \text{nm}$ with the cost of large side lobes in both the lateral and axial directions [Zhu20].

Besides, MLAs also suffer from fundamental trade-offs among NAs, pitches and working distances. High NAs, small pitches and long working distances cannot be achieved simultaneously because the lens elements must have pitches larger than their apertures. The DLAs can solve such a limitation by the overlap of the apertures of their lens elements [Wu10, Hul12, Ste16]. In contrast to the traditional lens arrays, the field distributions of their lens elements are superimposed with each other. A single lens element can receive contributions from adjacent elements to have a larger effective aperture, which releases the restriction between NAs, pitches and working distances. Thus, they can

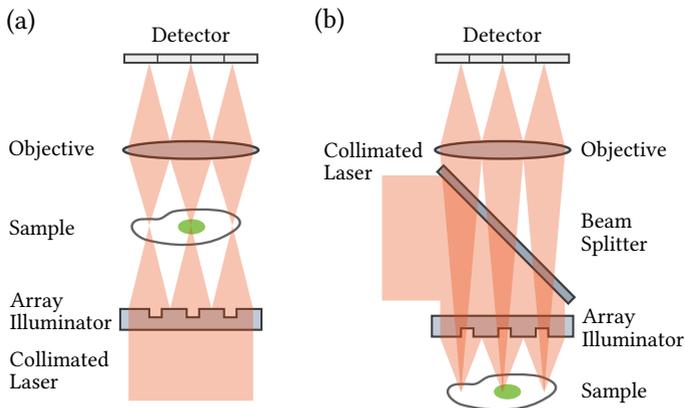


Figure 1.2: Schematics of microscope setups based on array illuminators. (a) A transmitted-light microscope configuration. (b) A reflected-light microscope configuration.

produce highly focused spots over large areas in dense grids. In confocal microscopes as Figure 1.2 (a) shows, the DLAs can utilize large FOVs of low-NA objectives, while the resolution can be enhanced by the tiny illumination spots produced by the DLAs. Such a principle is similar to that of the super-resolution microscopy [Hel94, Bet06]. In this way, the DLA-based confocal microscopes provide the possibilities to perform high-resolution large-area measurement. Furthermore, the DLAs can be easily replicated by lithography or compression molding at a low cost, which makes them promising for a wide range of applications.

In previous research, Wu et al. [Wu10] have shown a DLA producing a 200×40 spot array with an NA of 0.36, a pitch of $30 \mu\text{m}$ and a working distance of 6 mm. They have measured a spatial frequency up to $228 \text{ lp}\cdot\text{mm}^{-1}$ at a wavelength of 532 nm with an FOV of $6 \text{ mm} \times 5 \text{ mm}$ by a doublet lens pair as the objective. The DLA has been made by recording the interference pattern of a pinhole grid on a holographic plate. Hulsken et al. [Hul12] have demonstrated a DLA which has been designed by the angular spectrum method (ASM) with periodicity. It produces a spot array with an NA of 0.69, a working distance of $458 \mu\text{m}$ and a pitch of $15 \mu\text{m}$ at a wavelength of 650 nm. Scanning of kidney tissues has been demonstrated with the DLA. Stenau et al. [Ste16, Ste17] have made a DLA also by the ASM with a pitch of $44 \mu\text{m}$ and a working distance of 2 mm. The designed NA is 0.75, while the measured full width at half maximum (FWHM) of the spot is $0.443 \mu\text{m}$, which corresponds to an NA of 0.62. They have measured a spatial cut-off frequency of $900 \text{ lp}\cdot\text{mm}^{-1}$ at a wavelength of 532 nm by a 4×0.32 NA objective in the experiment.

However, a thorough search of the relevant literature shows that the previous works [Wu10, Hul12, Liu12, Liu14, Ste16, Ste17] only demonstrate the DLAs in transmitted-light microscopes similar to the one in Figure 1.2 (a), which can only be used for inspection of semi-transparent samples and are not preferable for surface metrology or fluorescence microscopy [Plo99, Rus09]. Using them in reflected-light microscopes with epi-illumination like the one in Figure 1.2 (b) will suffer from strong disturbance [Li19b], because the projected spots have to pass through the DLAs once again before entering the image system, which leads to a significant reduction of the signal-to-noise ratio (SNR) in the

images. Besides, previous DLAs designed by the ASM [Hul12, Liu12, Liu14, Ste16, Ste17] have infinitely extended periodic structures. For a finite grid in reality, the spots away from the center of the array become larger and fainter. Controlling the brightness of individual spots and further simulation of the imaging system are difficult by such a method. Moreover, there are also ambiguity problems with the spots, since there is no clear boundaries of the periodic array.

1.3 Main contributions

To solve the problems of the previous DLAs, new DLA concepts are proposed in this thesis. They utilize the principle of superposition of different field distributions to allow more flexible functionalities. They are successfully applied in reflected-light microscopes for measurement of opaque surfaces and fluorescent samples.

The See-through DLAs are proposed. The idea is to increase the zero-order diffraction efficiency of the DLAs by adding a plane wave component to the field distribution. In this way, the DLAs equivalently become more transparent and the SNR of the images can be significantly increased in the reflected-light microscope configurations [Li20b]. Two DLA prototypes with designed wavelengths of 785 nm and 488 nm have been produced. Experiments are carried out to verify their performance [Li20a].

- The diffraction simulation framework by the Rayleigh-Sommerfeld integral (RSI) is established to avoid periodicity and to better fulfill the sampling criteria, which also makes image simulation more flexible. A physically based model is proposed for the design and optimization of the DLAs to optimize the image SNR in the confocal setup, which has much fewer variables than traditional pixel-wise optimization methods [Li19a, Li19b].
- At 785 nm wavelength, the DLA produces an 11×11 spot array with a pitch of $100 \mu\text{m}$ and a working distance of $1095 \mu\text{m}$. The central spot has a lateral FWHM of $0.496 \mu\text{m}$, which corresponds to an NA of 0.78. A lateral spatial

cut-off frequency of $1024 \text{ lp}\cdot\text{mm}^{-1}$ has been measured by a $5\times 0.15 \text{ NA}$ objective in a reflected-light microscope configuration [Li21d].

- At 488 nm wavelength, the DLA produces an 13×13 spot array with a pitch of $80 \mu\text{m}$ and a working distance of $501 \mu\text{m}$. The central spot has a lateral FWHM of $0.304 \mu\text{m}$, which corresponds to an NA of 0.83. A lateral spatial cut-off frequency of $2048 \text{ lp}\cdot\text{mm}^{-1}$ has been measured by a $10\times 0.3 \text{ NA}$ objective, which shows the highest spatial resolution so far among the array illuminators to the best of found knowledge.
- The See-through DLA is also successfully applied in fluorescence microscopy [Li21e], which produces a 41×41 spot array with a pitch of $75 \mu\text{m}$, a working distance of $1001 \mu\text{m}$ and an NA of 0.83 at a wavelength of 488 nm. By measurement of the fluorescence beads, the setup shows the ability of high-resolution measurement over an area of $3 \text{ mm} \times 3 \text{ mm}$ with a $2.5\times 0.07 \text{ NA}$ objective, which results in an SBP of around 300 megapixels according the Rayleigh criterion [Li22]. Besides, the sample has been successfully measured by a simple doublet with strong aberrations without significant loss in resolution, which shows that the resolution of the setup is insensitive to optical aberrations.

The See-through DLAs demonstrate high lateral resolution in the experiments. However, when opaque surfaces are measured, the axial measurement sensitivity is still dominated by the objectives in theory. In order to increase the axial measurement sensitivity, various ideas are proposed.

- The Direct-imaging DLAs are proposed. They are designed by the superposition of two lens elements and act exactly as high-NA finite-conjugate objectives which can perform three-dimensional (3D) multi-spot confocal surface measurement with high resolution. Laterally, measurement of a resolution target shows a cut-off frequency of $1448 \text{ lp}\cdot\text{mm}^{-1}$ with an FOV of $0.5 \text{ mm} \times 0.5 \text{ mm}$ by a $5\times 0.15 \text{ NA}$ objective. Axially, an average height of 917.5 nm with a standard deviation of 49.9 nm is measured with a calibrated step height target which has a nominal height of 925.5 nm at a wavelength of 785 nm [Li20c, Li21c].

- The axial measurement sensitivity of the See-through DLAs can be further enhanced by interference between the probe wave and the reflected conjugate wave by the phase mask. An average height of 904.7 nm with a standard deviation of 9.7 nm is measured with a calibrated step height target at a wavelength of 785 nm [Li21a, Li21b].

Overall, various microscope concepts based on the DLAs are proposed in the thesis. Their capabilities are demonstrated both in theory and by experiments which are superior to that of the previous array illuminators in high-resolution large-area measurement. They show great potential in meeting the demands from industry and research in different applications like surface metrology and fluorescence microscopy.

1.4 Thesis outline

This work is organized in five chapters. Chapter 2 introduces the theoretical background for confocal imaging and diffraction simulation with scalar approximation. Numerical calculation, especially the sampling criteria, is extensively discussed to compare the accuracy of different diffraction propagation methods. Meanwhile, the theoretical basis for resolution enhancement by the DLAs in the confocal microscopes is derived from the classic imaging theory for scanning microscopes.

Chapter 3 introduces in detail the design and optimization processes of the See-through DLAs and the Direct-imaging DLAs. Both of them utilize superposition of different field components. The signal-to-background ratio of the spots in the images is used as the merit for optimization.

Chapter 4 presents the experiment results for characterization of the proposed DLAs. Different experiments including characterization of the spot sizes, lateral measurement of the resolution target, axial measurement of the step height target and imaging of the fluorescent samples are carried out to validate the performance of the proposed DLA-based microscope setups.

At last, Chapter 5 summarizes the achievement of this work. An outlook is also presented for possible improvements and further developments on this topic in the future.

2 Theoretical Background

This chapter introduces the theoretical background of the diffraction simulation and confocal imaging. In Section 2.1, the fundamentals of electrodynamics and Fourier optics are presented based on the classic books by Goodman [Goo05], Born and Wolf [Bor13], and the lecture notes of Theoretical Optics by Prof. Rockstuhl at Karlsruhe Institute of Technology (KIT). The introduction and derivation of the confocal imaging theory in Section 2.2 is mainly based on the book by Wilson and Sheppard [Wil84].

2.1 Diffraction simulation

Diffraction, as defined by Sommerfeld, is “any deviation of light rays from rectilinear paths which cannot be interpreted as reflection or refraction [Som54].” The diffraction theory is of major importance in optics and it also lays the foundation for this work. In this section, the basics of the diffraction simulation is presented. First, in Section 2.1.1, Maxwell’s equations are presented in different forms. The wave equation is derived under certain conditions. Then, in Section 2.1.2, the scalar approximation of the Helmholtz equation is introduced. Afterwards, several solutions to the wave equation are presented and discussed. At last, in Section 2.1.3, the numerical implementation of the solutions are introduced. Their sampling criteria and accuracy are compared in the scale of the applications in this work.

2.1.1 Maxwell's equations

The propagation of light is governed by Maxwell's equations [Max65]. The equations can be solved under certain initial and boundary conditions to describe different phenomena of light. The macroscopic form of Maxwell's equations in the time domain is presented as follows

$$\left\{ \begin{array}{l} \nabla \cdot \mathbf{D}(\mathbf{r},t) = \rho_f(\mathbf{r},t), \\ \nabla \cdot \mathbf{B}(\mathbf{r},t) = 0, \\ \nabla \times \mathbf{E}(\mathbf{r},t) = -\frac{\partial \mathbf{B}(\mathbf{r},t)}{\partial t}, \\ \nabla \times \mathbf{H}(\mathbf{r},t) = \mathbf{j}_f(\mathbf{r},t) + \frac{\partial \mathbf{D}(\mathbf{r},t)}{\partial t}, \end{array} \right. \quad (2.1)$$

where $\nabla \cdot$ is the divergence operator, $\nabla \times$ is the curl operator, $\mathbf{r} = (x, y, z)$ is the coordinate, t is the time, $\mathbf{D}(\mathbf{r},t)$ is the electric displacement in the units of $\text{A}\cdot\text{s}\cdot\text{m}^{-2}$, $\mathbf{B}(\mathbf{r},t)$ is the magnetic induction in the units of $\text{V}\cdot\text{s}\cdot\text{m}^{-2}$, $\mathbf{E}(\mathbf{r},t)$ is the electric field in the units of $\text{V}\cdot\text{m}^{-1}$, $\mathbf{H}(\mathbf{r},t)$ is the magnetic field strength in the units of $\text{A}\cdot\text{m}^{-1}$, $\rho_f(\mathbf{r},t)$ is the free charge density in the units of $\text{A}\cdot\text{s}\cdot\text{m}^{-3}$ and $\mathbf{j}_f(\mathbf{r},t)$ is the free current density in the units of $\text{A}\cdot\text{m}^{-2}$.

The electric field and magnetic induction are connected with their auxiliary fields according to the material properties. Such relations are known as the constitutive relations which are expressed as follows

$$\mathbf{D}(\mathbf{r},t) = \epsilon_0 \mathbf{E}(\mathbf{r},t) + \mathbf{P}(\mathbf{r},t), \quad (2.2)$$

$$\mathbf{H}(\mathbf{r},t) = \frac{1}{\mu_0} [\mathbf{B}(\mathbf{r},t) - \mathbf{M}(\mathbf{r},t)], \quad (2.3)$$

where $\mathbf{P}(\mathbf{r},t)$ is the polarization field in the units of $\text{A}\cdot\text{s}\cdot\text{m}^{-2}$, $\mathbf{M}(\mathbf{r},t)$ is the magnetization field in the units of $\text{V}\cdot\text{s}\cdot\text{m}^{-2}$, $\epsilon_0 \approx 8.854 \times 10^{-12} \text{ A}\cdot\text{s}\cdot\text{V}^{-1}\cdot\text{m}^{-1}$ is the vacuum permittivity, which is also known as the electric constant, and $\mu_0 \approx 1.257 \times 10^{-6} \text{ V}\cdot\text{s}\cdot\text{A}^{-1}\cdot\text{m}^{-1}$ is the vacuum permeability, which is also known as the magnetic constant.

The constitutive relations are the dependencies of the polarization and magnetization on the electric and magnetic fields. The relations describe the impact of materials to the electromagnetic fields. In optics, usually the materials are non-magnetic materials without free charges, in which the magnetization and free charge density are assumed to be zero

$$\mathbf{M}(\mathbf{r},t) = 0 \quad (2.4)$$

$$\rho_f(\mathbf{r},t) = 0. \quad (2.5)$$

Moreover, the macroscopic free current density is composed of the convection and the conduction current densities

$$\mathbf{j}_f(\mathbf{r},t) = \rho_f(\mathbf{r},t)\mathbf{v}_\rho + \mathbf{j}(\mathbf{r},t). \quad (2.6)$$

The convection current density $\rho_f(\mathbf{r},t)\mathbf{v}_\rho$ is caused by the movement of free volumetric charges, where \mathbf{v}_ρ is the velocity and ρ_f is zero under the above assumption. The conduction current density $\mathbf{j}(\mathbf{r},t)$ is caused by conductivity of the material and it is a function of the electric field. Since the convection current density is zero for the electrically neutral materials, the free current density can be replaced solely by the conduction current density. Therefore, by substituting Equations 2.2 to 2.6 into Equations 2.1, one can transform Maxwell's equations into the following form

$$\left\{ \begin{array}{l} \varepsilon_0 \nabla \cdot \mathbf{E}(\mathbf{r},t) = -\nabla \cdot \mathbf{P}(\mathbf{r},t), \\ \nabla \cdot \mathbf{H}(\mathbf{r},t) = 0, \\ \nabla \times \mathbf{E}(\mathbf{r},t) = -\mu_0 \frac{\partial \mathbf{H}(\mathbf{r},t)}{\partial t}, \\ \nabla \times \mathbf{H}(\mathbf{r},t) = \mathbf{j}(\mathbf{r},t) + \frac{\partial \mathbf{P}(\mathbf{r},t)}{\partial t} + \varepsilon_0 \frac{\partial \mathbf{E}(\mathbf{r},t)}{\partial t}. \end{array} \right. \quad (2.7)$$

The above time-domain fields can be transformed into the frequency domain by the Fourier transform. For example, for the electric field, it can be represented

by the inverse Fourier transform as

$$\mathbf{E}(\mathbf{r}, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \underline{\mathbf{E}}(\mathbf{r}, \omega) e^{-i\omega t} d\omega, \quad (2.8)$$

which means that an arbitrary field in the time domain can be treated as a composition of infinite series of time harmonic waves, where ω is the angular frequency of the waves and $\underline{\mathbf{E}}(\mathbf{r}, \omega)$ is the electric field strength in the frequency domain. Note that the Fourier transform may have different formulation concerning the prefactor and the sign of the exponential term, which does not affect the mathematical derivation or physical nature. In this thesis, the prefactor here is chosen to be $1/2\pi$ in the inverse transform, and thus the prefactor is unity in the forward transform. Similarly, the same transform can be applied to other fields in Equations 2.7, and the operator of partial derivative to time $\partial/\partial t$ is transformed into the multiplication with $-i\omega$ in the frequency domain. In this way, Maxwell's equations in the frequency domain can be obtained in the following form

$$\left\{ \begin{array}{l} \epsilon_0 \nabla \cdot \underline{\mathbf{E}}(\mathbf{r}, \omega) = -\nabla \cdot \underline{\mathbf{P}}(\mathbf{r}, \omega), \\ \nabla \cdot \underline{\mathbf{H}}(\mathbf{r}, \omega) = 0, \\ \nabla \times \underline{\mathbf{E}}(\mathbf{r}, \omega) = i\omega \mu_0 \underline{\mathbf{H}}(\mathbf{r}, \omega) \\ \nabla \times \underline{\mathbf{H}}(\mathbf{r}, \omega) = \underline{\mathbf{j}}(\mathbf{r}, \omega) - i\omega \underline{\mathbf{P}}(\mathbf{r}, \omega) - i\omega \epsilon_0 \underline{\mathbf{E}}(\mathbf{r}, \omega). \end{array} \right. \quad (2.9)$$

In the frequency domain, the response of the polarization field and current density to the electric field can be conveniently modeled as multiplication with a transfer function. The materials are restricted to be linear, homogeneous and isotropic. Thus, the relations are simplified and expressed as

$$\underline{\mathbf{P}}(\mathbf{r}, \omega) = \epsilon_0 \chi(\omega) \underline{\mathbf{E}}(\mathbf{r}, \omega), \quad (2.10)$$

$$\underline{\mathbf{j}}(\mathbf{r}, \omega) = \sigma(\omega) \underline{\mathbf{E}}(\mathbf{r}, \omega), \quad (2.11)$$

where $\chi(\omega)$ is the electric susceptibility and $\sigma(\omega)$ is the electric conductivity. By substituting Equations 2.10 and 2.11 into Equations 2.9, one can derive

the following expression as

$$\nabla \times [\nabla \times \underline{\mathbf{E}}(\mathbf{r}, \omega)] = i\omega\mu_0 \nabla \times \underline{\mathbf{H}}(\mathbf{r}, \omega) = \frac{\omega^2}{c_0^2} \left[1 + \chi(\omega) + \frac{i}{\omega\epsilon_0} \sigma(\omega) \right] \underline{\mathbf{E}}(\mathbf{r}, \omega), \quad (2.12)$$

where c_0 is the speed of light in vacuum. For the curl operator, there exists an identity as follows

$$\nabla \times [\nabla \times \underline{\mathbf{E}}(\mathbf{r}, \omega)] = \nabla[\nabla \cdot \underline{\mathbf{E}}(\mathbf{r}, \omega)] - \nabla^2 \underline{\mathbf{E}}(\mathbf{r}, \omega), \quad (2.13)$$

where ∇ is the gradient operator and ∇^2 is the Laplace operator. Besides, from the first equation in Equations 2.9 and Equation 2.10, the following relation can be obtained as

$$\epsilon_0[1 + \chi(\omega)]\nabla \cdot \underline{\mathbf{E}}(\mathbf{r}, \omega) = 0, \quad (2.14)$$

so the divergence of the electric field is zero. Thus, by putting Equation 2.12, 2.13 and 2.14 together, one can derive the wave equation in the frequency domain as

$$\nabla^2 \underline{\mathbf{E}}(\mathbf{r}, \omega) + \frac{\omega^2}{c_0^2} \left[1 + \chi(\omega) + \frac{i}{\omega\epsilon_0} \sigma(\omega) \right] \underline{\mathbf{E}}(\mathbf{r}, \omega) = 0. \quad (2.15)$$

Base on Equation 2.15, a complex dielectric function can be defined as

$$\epsilon(\omega) = 1 + \chi(\omega) + \frac{i}{\omega\epsilon_0} \sigma(\omega). \quad (2.16)$$

Together, they lead to the wave equation for the electric field, which is expressed as

$$\nabla^2 \underline{\mathbf{E}}(\mathbf{r}, \omega) + \frac{\omega^2}{c_0^2} \epsilon(\omega) \underline{\mathbf{E}}(\mathbf{r}, \omega) = 0. \quad (2.17)$$

Furthermore, by application of the same derivation to the \mathbf{H} -field, it is straightforward that it satisfies the identical equation, which is shown as

$$\nabla^2 \underline{\mathbf{H}}(\mathbf{r}, \omega) + \frac{\omega^2}{c_0^2} \epsilon(\omega) \underline{\mathbf{H}}(\mathbf{r}, \omega) = 0. \quad (2.18)$$

2.1.2 Scalar approximation and solutions

Equations 2.17 and 2.18 describe the propagation of light in non-magnetic, electrically neutral, linear, homogeneous and isotropic materials. Although it seems that a number of restrictions are applied, such materials are the most commonly used ones in optics, for example, various optical glasses.

The vectorial essence of the wave equation implies the coupling between different components of the electromagnetic field. Such coupling are usually weak or only limited to a certain region. For example, for diffraction of an aperture, the \mathbf{E} and \mathbf{H} fields only vary at the aperture edges where the material has interaction with light, which only affects the region of several wavelengths around the aperture [Goo05]. Usually, the aperture is much larger compared to the wavelength of light. Besides, it has been shown above that all components of \mathbf{E} and \mathbf{H} fields satisfy the wave equation in the same form. Thus, it is a natural choice to approximate the vectorial equation and to simplify the propagation of light in a scalar form, which is shown as

$$\nabla^2 u(\mathbf{r}, \omega) + \frac{\omega^2}{c_0^2} \epsilon(\omega) u(\mathbf{r}, \omega) = 0, \quad (2.19)$$

where $u(\mathbf{r}, \omega)$ represents the scalar field. Taking the inverse Fourier transform of the field into the spatial frequency domain can lead to the following expression

$$u(\mathbf{r}, \omega) = \frac{1}{(2\pi)^3} \int_{-\infty}^{\infty} U(\mathbf{k}, \omega) e^{i\mathbf{k} \cdot \mathbf{r}} dk_x dk_y dk_z, \quad (2.20)$$

where $\mathbf{k} = (k_x, k_y, k_z)$ is the wave vector. Note that in Equation 2.20, the phase factor in the integral $e^{i\mathbf{k}\cdot\mathbf{r}}$ has an opposite sign compared to the phase factor $e^{-i\omega t}$ in the inverse Fourier transform to time in Equation 2.8, because the oscillation factor of a forward propagating harmonic wave is defined as $e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}$. By taking the above expression into Equation 2.19, one can derive the following equation as

$$U(\mathbf{k}, \omega) \left[-\mathbf{k}^2 + \frac{\omega^2}{c_0^2} \varepsilon(\omega) \right] = 0, \quad (2.21)$$

from which the relation can be derived as

$$\mathbf{k}^2 = k(\omega)^2 = \frac{\omega^2}{c_0^2} \varepsilon(\omega), \quad (2.22)$$

where the scalar $k(\omega)$ is defined as the wave number. For simplicity, in the following derivations, the explicit dependency on ω is omitted and hence Equation 2.19 can be written as

$$\nabla^2 u(\mathbf{r}) + k^2 u(\mathbf{r}) = 0. \quad (2.23)$$

The mathematical form of Equation 2.23 is known as the scalar Helmholtz equation. It is an exact description for one-dimensional electromagnetic fields with a linear polarization, and it is an approximation in two-dimensional cases since it doesn't account for the coupling between different field components. However, there is a high computational cost for the vectorial equations. Although there have been enormous developments in modern computers and electromagnetic simulation algorithms, the vectorial methods are still too slow for large-scale simulations. For example, recent GPU-accelerated finite-element time-domain (FDTD) method can reach the calculation speed of around 10^9 cells per second [War19]. For a cubic volume of 1 mm^3 and a wavelength of 500 nm with a cell size of $\sqrt{3}\lambda/2$ and a time step of $\lambda/(2c_0)$, it still takes around 10 hours to complete the simulation for wave propagation by satisfying the Courant–Friedrichs–Lewy stability condition [Taf05].

In contrast, the scalar diffraction theory has the major advantages in simplicity and computation speed, and it is still being used to simulate a wide variety of diffraction phenomena with success. Its solutions have been extensively studied in comparison with the exact solutions to the vectorial equations. Generally, for DOEs, it is found that the scalar diffraction theory lead to a larger error with a small feature size, i.e., several wavelengths, or a large angle of incidence [Pom94, Ben98]. Nevertheless, the accuracy varies from different situations and highly depends on specific applications in practice [Mel01].

For the DLAs, it has been found that the scalar approximation still produces similar results compared to the rigorous coupled-wave analysis, which is an exact solution to Maxwell's equations, with a sub-wavelength feature size and an NA up to 0.9 [Ste17]. In the followings, different solutions to the scalar Helmholtz equation are derived. They have different preconditions and complexities, and therefore they are suitable for different application scenarios.

2.1.2.1 Rayleigh-Sommerfeld integral

Following the works of Huygens [Huy90] and Fresnel [Fre19], which explain diffraction in a conceptual way, Kirchoff pioneered in deriving a theory for light propagation directly from the wave equation [Kir83, Buc16]. As Figure 2.1 shows, the field $u(\mathbf{r})$ at any point \mathbf{r} behind a screen can be expressed by

$$u(\mathbf{r}) = \frac{1}{4\pi} \iint_{\Sigma + \bar{\Sigma}} G(\mathbf{r}, \mathbf{r}') \frac{\partial u(\mathbf{r}')}{\partial \mathbf{n}} - u(\mathbf{r}') \frac{\partial G(\mathbf{r}, \mathbf{r}')}{\partial \mathbf{n}} d^2 r', \quad (2.24)$$

where \mathbf{r}' is the coordinate on the screen, Σ is the surface on the aperture, $\bar{\Sigma}$ is the surface on the screen outside the aperture, the operator $\partial/\partial \mathbf{n}$ denotes the derivative along the outward normal direction to the boundary, and $G(\mathbf{r}, \mathbf{r}')$ is the Green's function of the Helmholtz equation, which satisfies

$$\nabla^2 G(\mathbf{r}, \mathbf{r}') + k^2 G(\mathbf{r}, \mathbf{r}') = \delta(\mathbf{r} - \mathbf{r}'), \quad (2.25)$$

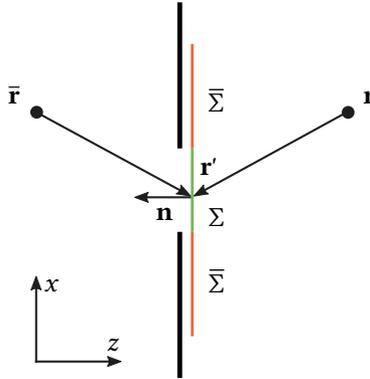


Figure 2.1: Sommerfeld's scheme of diffraction of an aperture on a plane screen.

while it is known to be

$$G(\mathbf{r}, \mathbf{r}') = \frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|}. \quad (2.26)$$

Equation 2.24 is valid under the Sommerfeld radiation condition for the boundary at infinity, which is shown as

$$\lim_{R \rightarrow \infty} \left[\frac{\partial u(\mathbf{r}')}{\partial \mathbf{n}} - iku(\mathbf{r}') \right] R = 0, \quad (2.27)$$

where R is the distance to the boundary. Kirchhoff then made further assumptions on the boundary conditions of the screen, where $u(\mathbf{r}')$ and its derivative $\partial u(\mathbf{r}')/\partial \mathbf{n}$ are zero across $\bar{\Sigma}$, which is known as the Kirchhoff boundary condition. Then Equation 2.24 simply becomes

$$u(\mathbf{r}) = \frac{1}{4\pi} \iint_{\underline{\Sigma}} G(\mathbf{r}, \mathbf{r}') \frac{\partial u(\mathbf{r}')}{\partial \mathbf{n}} - u(\mathbf{r}') \frac{\partial G(\mathbf{r}, \mathbf{r}')}{\partial \mathbf{n}} d^2 r'. \quad (2.28)$$

The above equation is the Kirchhoff's diffraction formula. Its detailed derivation can be found in Appendix A. However, although Kirchhoff's diffraction

formula produces peculiarly accurate results in experiments, its boundary condition is mathematically inconsistent. As was first pointed out by Poincaré [Poi89, Buc16], if both the field and its normal derivative is zero at a finite area, they must vanish everywhere else in the space as well.

Sommerfeld addressed the problem and derived the first exact solutions to the Helmholtz equation for diffraction under boundary conditions which are mathematically consistent [Som96, Som04, Som54], following Rayleigh's work of deriving a duo of solutions with Kirchhoff's requirement for $u(\mathbf{r}')$ and $\partial u(\mathbf{r}')/\partial \mathbf{n}$ separately [FRS97]. Consequently, these solutions are known as the Rayleigh-Sommerfeld integrals (RSIs).

In order to eliminate the integral on $\bar{\Sigma}$ in Equation 2.24, instead of restricting $u(\mathbf{r}')$ and $\partial u(\mathbf{r}')/\partial \mathbf{n}$ to be zero simultaneously, Sommerfeld chose a second Green's function which is generated from a mirrored point of \mathbf{r} with $\bar{\mathbf{r}} = (x, y, -z)$ as shown in Figure 2.1. Together, they can form the following functions

$$G_+(\mathbf{r}, \mathbf{r}') = \frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} + \frac{e^{ik|\bar{\mathbf{r}}-\mathbf{r}'|}}{|\bar{\mathbf{r}}-\mathbf{r}'|}, \quad (2.29)$$

$$G_-(\mathbf{r}, \mathbf{r}') = \frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} - \frac{e^{ik|\bar{\mathbf{r}}-\mathbf{r}'|}}{|\bar{\mathbf{r}}-\mathbf{r}'|}. \quad (2.30)$$

It is obvious that both $G_+(\mathbf{r}, \mathbf{r}')$ and $G_-(\mathbf{r}, \mathbf{r}')$ satisfy Equation 2.25. Therefore, they are also the Green's functions to the Helmholtz equation in the interested volume. Besides, on the plane screen which consists of Σ and $\bar{\Sigma}$, it is easy to know that $G_-(\mathbf{r}, \mathbf{r}') = 0$ and $\partial G_+(\mathbf{r}, \mathbf{r}')/\partial \mathbf{n} = 0$. Subsequently, by substituting $G_-(\mathbf{r}, \mathbf{r}')$ into Equation 2.24 and assuming $u(\mathbf{r}') = 0$ on $\bar{\Sigma}$ as the boundary condition, one can derive the first RSI as

$$u(\mathbf{r}) = -\frac{1}{4\pi} \iint_{\Sigma} u(\mathbf{r}') \frac{\partial G_-(\mathbf{r}, \mathbf{r}')}{\partial \mathbf{n}} d^2 r'. \quad (2.31)$$

Specifically, the derivative of the Green's function can be calculated as

$$\begin{aligned} \frac{\partial G_-(\mathbf{r}, \mathbf{r}')}{\partial \mathbf{n}} &= \cos \angle(\mathbf{n}, \mathbf{r} - \mathbf{r}') \left(ik - \frac{1}{|\mathbf{r} - \mathbf{r}'|} \right) \frac{e^{ik|\mathbf{r} - \mathbf{r}'|}}{|\mathbf{r} - \mathbf{r}'|} \\ &\quad - \cos \angle(\mathbf{n}, \bar{\mathbf{r}} - \mathbf{r}') \left(ik - \frac{1}{|\bar{\mathbf{r}} - \mathbf{r}'|} \right) \frac{e^{ik|\bar{\mathbf{r}} - \mathbf{r}'|}}{|\bar{\mathbf{r}} - \mathbf{r}'|}. \end{aligned} \quad (2.32)$$

From Figure 2.1, it is obvious that $\cos \angle(\mathbf{n}, \mathbf{r} - \mathbf{r}') = -\cos \angle(\mathbf{n}, \bar{\mathbf{r}} - \mathbf{r}')$ due to the symmetry, where \angle denotes the angle between the two vectors, and $|\mathbf{r} - \mathbf{r}'| = |\bar{\mathbf{r}} - \mathbf{r}'|$. Equation 2.32 can subsequently be simplified as

$$\frac{\partial G_-(\mathbf{r}, \mathbf{r}')}{\partial \mathbf{n}} = 2 \cos \angle(\mathbf{n}, \mathbf{r} - \mathbf{r}') \left(ik - \frac{1}{|\mathbf{r} - \mathbf{r}'|} \right) \frac{e^{ik|\mathbf{r} - \mathbf{r}'|}}{|\mathbf{r} - \mathbf{r}'|}. \quad (2.33)$$

Then, the first RSI can be explicitly written as

$$u(\mathbf{r}) = -\frac{1}{2\pi} \iint_{\Sigma} u(\mathbf{r}') \cos \angle(\mathbf{n}, \mathbf{r} - \mathbf{r}') \left(ik - \frac{1}{|\mathbf{r} - \mathbf{r}'|} \right) \frac{e^{ik|\mathbf{r} - \mathbf{r}'|}}{|\mathbf{r} - \mathbf{r}'|} d^2 r'. \quad (2.34)$$

It is worth noting that the first RSI can be further simplified with the assumption $|\mathbf{r} - \mathbf{r}'| \gg \lambda$, and Equation 2.34 becomes

$$u(\mathbf{r}) = -\frac{ik}{2\pi} \iint_{\Sigma} u(\mathbf{r}') \cos \angle(\mathbf{n}, \mathbf{r} - \mathbf{r}') \frac{e^{ik|\mathbf{r} - \mathbf{r}'|}}{|\mathbf{r} - \mathbf{r}'|} d^2 r'. \quad (2.35)$$

Similarly, substituting $G_+(\mathbf{r}, \mathbf{r}')$ into Equation 2.24 and assuming the normal derivative of the field $\partial u(\mathbf{r}')/\partial \mathbf{n} = 0$ on $\bar{\Sigma}$ as the boundary condition, one can derive the second RSI as

$$u(\mathbf{r}) = \frac{1}{4\pi} \iint_{\Sigma} G_+(\mathbf{r}, \mathbf{r}') \frac{\partial u(\mathbf{r}')}{\partial \mathbf{n}} d^2 r'. \quad (2.36)$$

By the fact that $|\mathbf{r} - \mathbf{r}'| = |\bar{\mathbf{r}} - \mathbf{r}'|$, Equation 2.36 can be written as

$$u(\mathbf{r}) = \frac{1}{2\pi} \iint_{\Sigma} \frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r} - \mathbf{r}'|} \frac{\partial u(\mathbf{r}')}{\partial \mathbf{n}} d^2r'. \quad (2.37)$$

By comparison of the two RSIs with Kirchhoff's diffraction formula, it is found that the latter one is indeed an average of the first two integrals, since there are obvious relations among the Green's functions with $G_+(\mathbf{r}, \mathbf{r}') = 2G(\mathbf{r}, \mathbf{r}')$ and $\partial G_-(\mathbf{r}, \mathbf{r}')/\partial \mathbf{n} = 2\partial G(\mathbf{r}, \mathbf{r}')/\partial \mathbf{n}$. Besides, despite its mathematical inconsistency, Kirchhoff's formula can be applied to surfaces with any shapes, while the RSIs can only be applied on planar surfaces, as restricted by the Green's functions. Nevertheless, there are plenty of research to compare the three formulas, which have shown little differences for apertures much larger than the wavelength and distances far away from the apertures [Wol64, Heu73]. In reality, the first RSI is commonly used because of its simplicity compared to the other two solutions [Goo05].

2.1.2.2 Angular spectrum method

Another exact solution to the Helmholtz equation is derived by the angular spectrum method (ASM) [Boo50, Cle51], which is the decomposition of the electromagnetic fields into a series of plane waves. The method can calculate the field distribution at $z > 0$ in the propagation direction according to the initial field at $z = 0$, which is shown in Figure 2.2.

By the Fourier transform of the scalar field into spatial frequency domain in Equation 2.20 and the dispersion relation in Equation 2.22, the three components of the wave vector are restricted by

$$k_x^2 + k_y^2 + k_z^2 = k^2 = \frac{\omega^2}{c_0^2} \epsilon. \quad (2.38)$$

Note again that the frequency dependency of the wave number, etc., are hidden for simplicity. Accordingly, there are only two independent variables in the

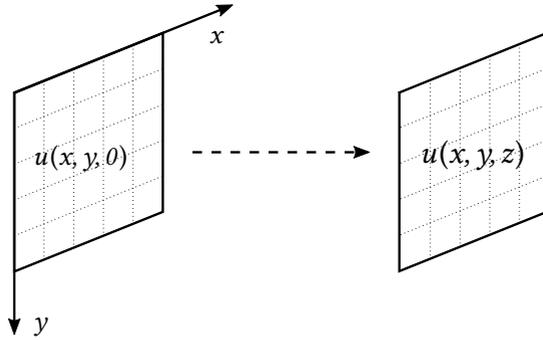


Figure 2.2: Angular spectrum propagation between two parallel planes.

spatial frequency, and the third one can be calculated by the other two with the above dispersion relation. For simplicity, the spatial frequency components are replaced by Greek letters as

$$k_x = \alpha, \quad k_y = \beta, \quad k_z = \gamma. \quad (2.39)$$

Then, a 2D inverse Fourier transform can be applied to the field as

$$u(\mathbf{r}) = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} U(\alpha, \beta, z) e^{i(\alpha x + \beta y)} d\alpha d\beta. \quad (2.40)$$

By substituting Equation 2.40 into the Helmholtz equation 2.23, one can simply derive that

$$\left(\frac{d^2}{dz^2} - \alpha^2 - \beta^2 \right) U(\alpha, \beta, z) e^{i(\alpha x + \beta y)} + k^2 U(\alpha, \beta, z) e^{i(\alpha x + \beta y)} = 0. \quad (2.41)$$

By expressing γ as a function of α and β according to Equation 2.38, one can make further simplification which leads to the following equation

$$\left(\frac{d^2}{dz^2} + \gamma(\alpha, \beta)^2 \right) U(\alpha, \beta, z) = 0. \quad (2.42)$$

The general solution to the above partial differential equation is straightforward, which is shown as

$$U(\alpha, \beta, z) = C_1 e^{i\gamma(\alpha, \beta)z} + C_2 e^{-i\gamma(\alpha, \beta)z}, \quad (2.43)$$

where C_1 and C_2 are arbitrary constants. According to the physical property of the wave vector, for $\gamma(\alpha, \beta)^2 > 0$, it indicates that the wave is propagating with a real wave vector, and for $\gamma(\alpha, \beta)^2 < 0$, it indicates that the wave is evanescent with an imaginary wave vector. For imaginary γ , the second term in Equation 2.43 will grow exponentially with the increase of z , which is physically unrealistic and therefore should be suppressed. Subsequently, with the boundary condition $U(\alpha, \beta, 0) = U_0(\alpha, \beta)$, Equation 2.43 becomes

$$U(\alpha, \beta, z) = U_0(\alpha, \beta) e^{i\gamma(\alpha, \beta)z}, \quad (2.44)$$

where $U_0(\alpha, \beta)$ is the Fourier transform of the field at $z = 0$, which is expressed as

$$U_0(\alpha, \beta) = \int_{-\infty}^{\infty} u(x, y, 0) e^{-i(\alpha x + \beta y)} dx dy. \quad (2.45)$$

As a result, Equation 2.40 can be transformed into

$$u(\mathbf{r}) = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} U_0(\alpha, \beta) e^{i\gamma(\alpha, \beta)z} e^{i(\alpha x + \beta y)} d\alpha d\beta, \quad (2.46)$$

which is the formula of ASM. Equation 2.46 indicates that the field at any position z after the diffraction propagation can be calculated by the initial field at $z = 0$ and the dispersion relation of the material, while the initial field $u(x, y, 0) = u_0(x, y)$ serves as the boundary condition. The ASM can be easily formulated by computer programs and it has high computation efficiency which benefits from the fast Fourier transform (FFT) algorithm. Hence, it is a commonly used method for diffraction simulation in research and commercial

software. However, it also demands high sampling rate when the propagation distance is relatively long, which will be discussed in detail in Section 2.1.3.

2.1.2.3 Fresnel approximation

Another widely used approach for diffraction simulation is the Fresnel diffraction integral. With the further assumption of $x, y \ll z$ and $x', y' \ll z$, which is known as the Fresnel approximation or the paraxial approximation, the components in the first RSI (Equation 2.35) can be further simplified.

First, the cosine of the angle between \mathbf{n} and $\mathbf{r} - \mathbf{r}'$ can be approximated by

$$\cos \angle(\mathbf{n}, \mathbf{r} - \mathbf{r}') \approx 1. \quad (2.47)$$

Second, the denominator in the fraction term which affects the amplitude of the wave can be safely simplified as

$$\frac{1}{|\mathbf{r} - \mathbf{r}'|} \approx \frac{1}{z}. \quad (2.48)$$

At last, the oscillating phase can be approximated by the Taylor series with

$$\begin{aligned} k|\mathbf{r} - \mathbf{r}'| &= k\sqrt{z^2 + (x - x')^2 + (y - y')^2} \\ &\approx kz \left[1 + \frac{(x - x')^2}{2z^2} + \frac{(y - y')^2}{2z^2} \right]. \end{aligned} \quad (2.49)$$

In consequence, by the above approximations, the first RSI in Equation 2.35 can be simplified as

$$u_{\text{Fresnel}}(\mathbf{r}) = \frac{ke^{ikz}}{i2\pi z} \iint_{\Sigma} u(\mathbf{r}') e^{\frac{ik}{2z}[(x-x')^2 + (y-y')^2]} d^2r'. \quad (2.50)$$

Assuming the aperture surface Σ is located at $u(x', y', 0) = u_0(x', y')$, one can reformulate the above equation as

$$u_{\text{Fresnel}}(\mathbf{r}) = \frac{ke^{ikz}e^{\frac{ik}{2z}(x^2+y^2)}}{i2\pi z} \iint_{\Sigma} \left[u_0(x', y')e^{\frac{ik}{2z}(x'^2+y'^2)} \right] e^{-\frac{ik}{z}(xx'+yy')} dx' dy'. \quad (2.51)$$

The advantage of Equation 2.51 thus becomes apparent. According to its mathematical structure, the diffracted field can be simply calculated by the Fourier transform of the initial field with an additional phase factor. The Fresnel diffraction integral is also widely applied, and it is accurate enough under the circumstances where the paraxial approximation is valid. It has the great advantage in computation speed since only one Fourier transform is required for the calculation. Therefore, the Fresnel diffraction formula is one of the most commonly used simulation methods for DOE design. It is convenient to combine it with various optimization algorithms to achieve the design target.

2.1.2.4 Fraunhofer approximation

The Fresnel approximation can be even further simplified with the assumption that

$$z \gg \frac{k(x'^2 + y'^2)}{2}, \quad (2.52)$$

which is known as the Fraunhofer or far-field approximation. With the above approximation, the term in Equation 2.49 which contains x'^2 and y'^2 can be neglected, and accordingly it becomes

$$k|\mathbf{r} - \mathbf{r}'| \approx kz \left[1 - \frac{xx' + yy'}{z^2} + \frac{x^2 + y^2}{2z^2} \right]. \quad (2.53)$$

With the above approximation, the Fresnel integral in Equation 2.51 can be simplified into the Fraunhofer integral as

$$\begin{aligned} u_{\text{Fraunhofer}}(\mathbf{r}) &= \frac{ke^{ikz}e^{\frac{ik}{2z}(x^2+y^2)}}{i2\pi z} \iint_{\Sigma} u_0(x', y') e^{-\frac{ik}{z}(xx'+yy')} dx' dy' \\ &= \frac{ke^{ikz}e^{\frac{ik}{2z}(x^2+y^2)}}{i2\pi z} U_0\left(\frac{k}{z}x, \frac{k}{z}y\right), \end{aligned} \quad (2.54)$$

where U_0 is the Fourier transform of the initial field, which is expressed as

$$U_0\left(\frac{k}{z}x, \frac{k}{z}y\right) = \iint_{\Sigma} u_0(x', y') e^{-\frac{ik}{z}(xx'+yy')} dx' dy'. \quad (2.55)$$

The Fraunhofer diffraction formula indicates that the far-field diffraction pattern is merely the Fourier transform of the initial field. The far-field propagation can be equivalently emulated by a collimating lens. It is a very useful property which is commonly used for filtering of spatial frequencies.

2.1.3 Numerical implementation

For the application in this work, it is important to choose a suitable diffraction simulation method for the design of the DLAs. In the previous subsections, different solutions are introduced for the calculation of diffraction propagation.

The Fresnel and Fraunhofer diffraction formulas are based on the paraxial or far-field approximations. They are usually the most widely used methods for the design of DOEs. However, the DLAs in this work aim to achieve considerably high NAs, i.e., greater than 0.7. Therefore, the assumptions behind these two methods are not valid anymore in such situations.

On the contrary, both the RSI and the ASM are exact solutions of the Helmholtz equation under the scalar approximation. Analytically, they should produce the same results with the same boundary conditions. However, the nature of discrete computation by computer simulation makes many differences between them. They have different sampling requirements and will produce different

results even with exactly the same boundary conditions. For this reason, it is necessary to further evaluate their numerical properties. In the next sections, the RSI and the ASM are discussed in detail in different aspects. Their corresponding numerical formulation are introduced. Their sampling criteria are discussed and their accuracy in the scale of this work are compared.

2.1.3.1 Numerical formulation of the RSI

Recall the first RSI in Equation 2.31. By restricting the diffraction propagation between two parallel planes, which is the same configuration as the ASM in Figure 2.2, one can fix the initial plane at $z' = 0$ and the output plane at z , while the derivative to the normal direction $\partial/\partial\mathbf{n}$ can consequently be simplified as $\partial/\partial z$. Besides, considering the fact that $\partial G_-(\mathbf{r}, \mathbf{r}')/\partial\mathbf{n} = 2\partial G(\mathbf{r}, \mathbf{r}')/\partial\mathbf{n}$ and substituting the Green's function, one can rewrite Equation 2.31 as

$$u(x, y, z) = -\frac{1}{2\pi} \iint_{-\infty}^{\infty} u(x', y', 0) \frac{\partial}{\partial z} \left[\frac{e^{ik\sqrt{(x-x')^2+(y-y')^2+z^2}}}{\sqrt{(x-x')^2+(y-y')^2+z^2}} \right] dx' dy'. \quad (2.56)$$

Note that the integral here can also be equivalently carried out over the whole screen at $z' = 0$ in Figure 2.1 instead of on the aperture Σ solely, since the fields only have non-zero values on the aperture, which is one of the boundary conditions of the first RSI. It can be easily pointed out that the integral in Equation 2.56 is a convolution of two functions, which can be written as

$$u_z(x, y) = u_0(x, y) * h_z(x, y), \quad (2.57)$$

where $*$ denotes the convolution operator, $u_0(x, y) = u(x, y, 0)$, $u_z(x, y) = u(x, y, z)$ and $h_z(x, y) = h(x, y, z)$ is the impulse response function of the RSI. It is expressed as

$$\begin{aligned} h_z(x, y) &= -\frac{1}{2\pi} \frac{\partial}{\partial z} \left(\frac{e^{ikr}}{r} \right) \\ &= \frac{1}{2\pi} \left(\frac{1}{r} - ik \right) \frac{ze^{ikr}}{r^2}, \end{aligned} \quad (2.58)$$

where $r = \sqrt{x^2 + y^2 + z^2}$. Therefore, Equation 2.57 shows that the integral can be directly calculated by convolution.

Numerically, suppose that the initial field $u_0(x, y)$ are discretized into an equidistant $M \times N$ grid with sampling intervals of Δx and Δy . The convolution can thus be calculated discretely as the Riemann sum as

$$u_z(x, y) = \sum_{p=0}^{M-1} \sum_{q=0}^{N-1} u_0(x'_p, y'_q) \cdot h_z(x - x'_p, y - y'_q) \Delta x \Delta y, \quad (2.59)$$

where $x'_p = p\Delta x$ and $y'_q = q\Delta y$. Equation 2.59 is the most straightforward numerical formulation for the RSI, which is also known as the direct integration method. It has the advantage of flexibility since the position of the output field can be arbitrarily chosen. However, its main drawback is the low efficiency of the discrete linear convolution, which has a computational complexity of $O(M^2N^2)$ if the target field $u_z(x, y)$ is also an $M \times N$ grid [She06].

On the other hand, numerical convolution can be accelerated by FFT to reduce the computational cost. Equation 2.57 can be transformed into the frequency domain by the well-known convolution theorem. Let the two-dimensional Fourier transform and the inverse Fourier transform be defined by the following operators $\mathcal{F}\{\cdot\}$ and $\mathcal{F}^{-1}\{\cdot\}$ respectively, which are shown as

$$U_z(\alpha, \beta) = \mathcal{F}\{u_z(x, y)\} = \iint_{-\infty}^{\infty} u_z(x, y) e^{-i(\alpha x + \beta y)} dx dy, \quad (2.60)$$

$$u_z(x, y) = \mathcal{F}^{-1}\{U_z(\alpha, \beta)\} = \frac{1}{4\pi^2} \iint_{-\infty}^{\infty} U_z(\alpha, \beta) e^{i(\alpha x + \beta y)} d\alpha d\beta. \quad (2.61)$$

Thereafter, by the convolution theorem, linear convolution in the spatial domain is equivalent to multiplication in the frequency domain, and vice versa. So Equation 2.57 can be transformed into

$$u_z(x, y) = u_0(x, y) * h_z(x, y) = \mathcal{F}^{-1}\{\mathcal{F}[u_0(x, y)] \cdot \mathcal{F}[h_z(x, y)]\}. \quad (2.62)$$

For the numerical calculation, the field on the 2D plane is discretized, and it can be denoted as a matrix, which is represented as

$$\mathbf{u}_z[m, n] = u_z(x_m, y_n) = u_z(m\Delta x, n\Delta y), \quad (2.63)$$

where m and n are integer indices. Suppose that the matrix has a size of $M \times N$. Accordingly, the Fourier transform can be imitated by the discrete Fourier transform (DFT), and Equations 2.60 and 2.61 become

$$\mathbf{U}[p, q] = \text{DFT} \{ \mathbf{u} \} [p, q] = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \mathbf{u}[m, n] e^{-i2\pi(\frac{mp}{M} + \frac{nq}{N})}, \quad (2.64)$$

$$\mathbf{u}[m, n] = \text{IDFT} \{ \mathbf{U} \} [m, n] = \frac{1}{MN} \sum_{p=0}^{M-1} \sum_{q=0}^{N-1} \mathbf{U}[p, q] e^{i2\pi(\frac{mp}{M} + \frac{nq}{N})}, \quad (2.65)$$

where p and q are also integer indices, and IDFT denotes the inverse discrete Fourier transform. The fields after DFT implicitly become periodic with a period of M in x direction and N in y direction for the $M \times N$ matrices, which can be shown as

$$\mathbf{U}[p + s_1M, q + s_2N] = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \mathbf{u}[m, n] e^{-i2\pi(\frac{mp}{M} + \frac{nq}{N} + s_1m + s_2n)} = \mathbf{U}[p, q], \quad (2.66)$$

$$\mathbf{u}[m + s_1M, n + s_2N] = \frac{1}{MN} \sum_{p=0}^{M-1} \sum_{q=0}^{N-1} \mathbf{U}[p, q] e^{i2\pi(\frac{mp}{M} + \frac{nq}{N} + s_1p + s_2q)} = \mathbf{u}[m, n], \quad (2.67)$$

where s_1 and s_2 are arbitrary integers. Such periodicity also changes the convolution theorem in Equation 2.62, which makes the linear convolution into circular convolution [Opp09]. In practice, the DFT is usually calculated by the FFT, which is a family of algorithms producing the same results as the DFT with much higher computation efficiencies [Sch96]. So in the following text, the FFT will be used instead of the DFT. With the discretized field and

impulse response function, the circular convolution theorem is represented as

$$\text{IFFT} \{ \text{FFT} \{ \mathbf{u}_0 \} \odot \text{FFT} \{ \mathbf{h}_z \} \} = \mathbf{u}_0 \circledast \mathbf{h}_z, \quad (2.68)$$

where IFFT is the inverse fast Fourier transform, \odot denotes the element-wise multiplication and the operator \circledast denotes the circular convolution, which has the expression of

$$(\mathbf{u}_0 \circledast \mathbf{h}_z)[m,n] = \sum_{p=0}^{M-1} \sum_{q=0}^{N-1} \mathbf{u}_0[p,q] \cdot \mathbf{h}_z[\text{mod}(m-p, M), \text{mod}(n-q, N)], \quad (2.69)$$

where $\text{mod}(\cdot, N)$ refers to the modulo operation to N . The above expression is different than Equation 2.59, since the impulse response is used periodically in the calculation. On the other hand, to make it clearer, the circular operation is reversible, which is shown as

$$\mathbf{u}_0 \circledast \mathbf{h}_z = \mathbf{h}_z \circledast \mathbf{u}_0. \quad (2.70)$$

In consequence, it should be noted that the impulse response function is not convoluted with the initial field \mathbf{u}_0 on the aperture surrounded by 0, but the periodic version of it surrounded by its replicas. Thus, the RSI calculation in Equation 2.57 no longer holds for the discrete circular convolution. Direct replacement of the Fourier transform in Equation 2.62 by FFT causes errors, by which the resulting field will receive contributions from the virtual replicas of the initial field.

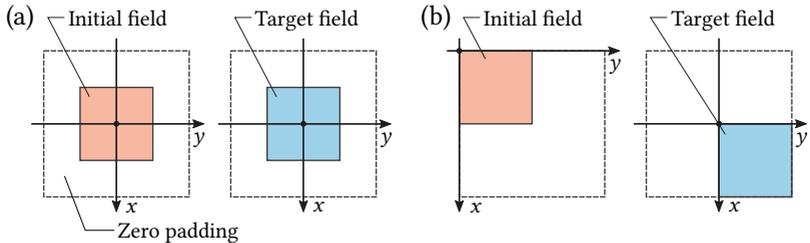


Figure 2.3: Calculation of linear convolution by FFT through zero-padding. (a) Field positioned in the center. (b) Field positioned at the corner. [Mat09]

Therefore, the correct way to calculate the linear convolution by FFT is to add zero-padding to the initial field $\mathbf{u}_0[m,n]$ [Mat09], which is shown by Figure 2.3. A new matrix is constructed and with double the size of the initial field. The initial field is positioned either in the center or at the corner of the padded matrix and it is surrounded by zero. In this way, the linear convolution of the initial field can be obtained either by FFT or circular convolution and the resulting field is positioned at the places as shown in Figure 2.3.

In order to calculate the RSI in Equation 2.59 by FFT to increase the speed, the equation can be expressed in the following way [She06]

$$\hat{\mathbf{u}}_z = \text{IFFT} \left\{ \text{FFT} \{ \hat{\mathbf{u}}_0 \} \odot \text{FFT} \{ \hat{\mathbf{h}}_z \} \right\} \Delta x \Delta y, \quad (2.71)$$

where the matrices are constructed as follows

$$\hat{\mathbf{u}}_0 = \begin{bmatrix} \mathbf{u}_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} = \left[\begin{array}{ccc|c} u_0(x_0, y_0) & \cdots & u_0(x_0, y_{N-1}) & \mathbf{0} \\ \vdots & \ddots & \vdots & \\ u_0(x_{M-1}, y_0) & \cdots & u_0(x_{M-1}, y_{N-1}) & \\ \hline & \mathbf{0} & & \mathbf{0} \end{array} \right]_{(2M-1) \times (2N-1)}, \quad (2.72)$$

$$\hat{\mathbf{h}}_z = \begin{bmatrix} h_z(x_{1-M}, y_{1-N}) & \cdots & h_z(x_{1-M}, y_{N-1}) \\ \vdots & \ddots & \vdots \\ h_z(x_{M-1}, y_{1-N}) & \cdots & h_z(x_{M-1}, y_{N-1}) \end{bmatrix}_{(2M-1) \times (2N-1)}. \quad (2.73)$$

The formula for the elements in $\hat{\mathbf{h}}_z$ and the notation for the coordinates can be found in Equations 2.58 and 2.63 respectively. The resulting target field can be obtained as shown in Figure 2.3 (b), by the following expression

$$\mathbf{u}_z[m, n] = \hat{\mathbf{u}}_z[m+M, n+N]. \quad (2.74)$$

The above numerical formulation is denoted as the FFT-RSI [She06]. It has the advantage over the direct integration of RSI in terms of the computation speed. Due to the use of FFT, the computational complexity is largely reduced to $O(4MN \log_2(4MN))$, which is much less than the $O(M^2N^2)$ of the summation

in Equation 2.59 especially for arrays with large sizes. According to the above equations, any fields after the propagation can be calculated by knowing the initial field, the propagation distance and the wave number in the medium.

2.1.3.2 Numerical formulation of the ASM

Recall the ASM formula in Equation 2.46. It can also be represented by the Fourier transform as the following equation shows

$$\begin{aligned} u_z(x, y) &= \mathcal{F}^{-1} \left\{ U_0(\alpha, \beta) \cdot e^{i\gamma(\alpha, \beta)z} \right\} \\ &= \mathcal{F}^{-1} \left\{ \mathcal{F} \left\{ u_0(x, y) \right\} \cdot H_z(\alpha, \beta) \right\}, \end{aligned} \quad (2.75)$$

where $H_z(\alpha, \beta)$ denotes the transfer function of the ASM, which has the expression of

$$H_z(\alpha, \beta) = e^{i\gamma(\alpha, \beta)z} = e^{i\sqrt{k^2 - \alpha^2 - \beta^2}z}. \quad (2.76)$$

Remarkably, by comparing the ASM in Equation 2.75 with the RSI Equation 2.62, one can find that they are inherently identical, provided the relation of

$$h_z(x, y) = \mathcal{F}^{-1} \left\{ H_z(\alpha, \beta) \right\}. \quad (2.77)$$

The above relation can be proved by the well-known Weyl identity [Wey19, Lal68], which has the expression of

$$\frac{e^{ikr}}{r} = \frac{i}{2\pi} \iint_{-\infty}^{\infty} \frac{e^{i\gamma(\alpha, \beta)|z|}}{\gamma(\alpha, \beta)} e^{i(\alpha x + \beta y)} d\alpha d\beta. \quad (2.78)$$

Assuming $z > 0$, by taking the derivative $\partial/\partial z$ and applying a prefactor of $-1/2\pi$ on both sides of the above identity, one can easily get

$$-\frac{1}{2\pi} \frac{\partial}{\partial z} \frac{e^{ikr}}{r} = \frac{1}{4\pi^2} \iint_{-\infty}^{\infty} e^{i\gamma(\alpha, \beta)z} e^{i(\alpha x + \beta y)} d\alpha d\beta, \quad (2.79)$$

which is exactly the relation in Equation 2.77.

Similarly, Equation 2.75 can be computed numerically by discretizing the field and the transfer function. Besides, the discretized matrices also need zero-padding in order to avoid the periodicity and overlapping caused by the circular convolution. In a way that is comparable to Equation 2.71, the numerical formula for the ASM is shown as the following expression

$$\hat{\mathbf{u}}_z = \text{IFFT} \left\{ \text{FFT} \{ \hat{\mathbf{u}}_0 \} \odot \hat{\mathbf{H}}_z \right\}, \quad (2.80)$$

where $\hat{\mathbf{u}}_0$ has the same expression as Equation 2.72, and the matrix for the transfer function is constructed as

$$\hat{\mathbf{H}}_z = \begin{bmatrix} H_z(\alpha_{1-M}, \beta_{1-N}) & \cdots & H_z(\alpha_{1-M}, \beta_{N-1}) \\ \vdots & \ddots & \vdots \\ H_z(\alpha_{M-1}, \beta_{1-N}) & \cdots & H_z(\alpha_{M-1}, \beta_{N-1}) \end{bmatrix}_{(2M-1) \times (2N-1)}. \quad (2.81)$$

The notations for the spatial frequencies α and β are shown by the following expressions

$$\alpha_p = \frac{2\pi p}{M} \cdot \frac{1}{2\Delta x}, \quad (2.82)$$

$$\beta_q = \frac{2\pi q}{N} \cdot \frac{1}{2\Delta y}. \quad (2.83)$$

At last, the target field \mathbf{u}_z can be extracted from the resulting zero-padded field $\hat{\mathbf{u}}_0$ according to Equation 2.74. The above numerical formulation for the ASM is denoted as the FFT-ASM. It has the same computational complexity as the FFT-RSI, while the FFT-RSI requires three FFT operations and the FFT-ASM only requires two FFT operations. Although the two methods are analytically identical, they have different sampling criteria and they may produce different numerical results under the same conditions. Their characteristics and differences are compared in the next sections.

2.1.3.3 Sampling criteria of the FFT-RSI and FFT-ASM

Before the detailed discussion, it is worth noting that the research of the scalar diffraction theory is still ongoing, especially on various numerical implementations mainly to save the computation cost [Meh17], for example, by altering the convolution kernel [Eng04, Rit14, Koz15], changing the sampling rates by non-uniform FFT [Shi12, Kim14, Wan19, Zha20a], or limiting the bandwidth of the object or the propagation kernel [Zha06, Mat09]. Some of these methods accelerate the computation at a cost of accuracy, and others require specific arrangement of the sampling grid. For the application in this work, design of the high-NA DLAs requires simulation with as few assumptions as possible, while reduced sampling at certain places also has no benefits for the production or the performance of the DLAs. Therefore, this work still focuses on the FFT-RSI and the FFT-ASM, which are among the most well-proven methods so far for diffraction simulation.

The sampling of both methods is associated with the initial field and the impulse response function for the FFT-RSI or the transfer function for the FFT-ASM. For the field distribution u_0 , its bandwidth is determined by itself. For some types of field distributions, for example, an airy disk, it is inherently band-limited. Then, the sampling needs to satisfy the well-known Nyquist–Shannon sampling theorem [Sha49], which is expressed as

$$f_s \geq 2f_{\max} \quad (2.84)$$

where f_s is the sampling frequency, and f_{\max} is the highest frequency in the signal. For other types of fields, for example, an aperture with sharp edges, its Fourier transform U_0 has to be truncated at a maximum frequency in numerical calculation. However, the effect of the frequency cut is usually negligible, as long as the frequency range with most of the energy is sufficiently sampled.

The major issue for the sampling of the numerical diffraction simulation lies in the impulse response function $h_z(x, y)$ for the FFT-RSI and the transfer function $H_z(\alpha, \beta)$ for the FFT-ASM. Recall their expressions in Equation 2.58 and Equation 2.76 respectively. By taking a closer look at the two equations, one can find that the modulation of each of the functions has a chirp-like phase

in the exponential terms, which requires to be well sampled. The oscillating phase can be extracted and expressed as

$$\varphi_{\text{RSI}}(x, y) = k\sqrt{x^2 + y^2 + z^2}, \quad (2.85)$$

$$\varphi_{\text{ASM}}(\alpha, \beta) = z\sqrt{k^2 - \alpha^2 - \beta^2}, \quad (2.86)$$

where $\varphi_{\text{RSI}}(x, y)$ is the oscillating phase for the FFT-RSI, and $\varphi_{\text{ASM}}(x, y)$ is the oscillating phase for the FFT-ASM. Without loss of generality, the sampling in the x direction is first evaluated. By the sampling theorem in the inequality 2.84, the following expressions have to be fulfilled for the FFT-RSI and the FFT-ASM respectively [Zha20b]

$$\frac{1}{\Delta x} \geq 2 \left| \frac{1}{2\pi} \frac{\partial \varphi_{\text{RSI}}(x, y)}{\partial x} \right|_{\max}, \quad (2.87)$$

$$\frac{1}{2\pi \Delta f_x} \geq 2 \left| \frac{1}{2\pi} \frac{\partial \varphi_{\text{ASM}}(\alpha, \beta)}{\partial \alpha} \right|_{\max}, \quad (2.88)$$

where Δx is the sampling interval in space, and $\Delta f_x = 1/(2M\Delta x)$ is the sampling interval in frequency domain. By substituting Equations 2.85 and 2.86 and calculating the derivatives in the above inequalities, one can get

$$\frac{1}{\Delta x} \geq \left| \frac{kx}{\pi \sqrt{x^2 + y^2 + z^2}} \right|_{\max}, \quad (2.89)$$

$$\frac{1}{2\pi \Delta f_x} \geq \left| \frac{z\alpha}{\pi \sqrt{k^2 - \alpha^2 - \beta^2}} \right|_{\max}. \quad (2.90)$$

From the above expressions, it can be found that the sampling criteria of the two methods are different. For the FFT-RSI, apparently a smaller interval, i.e., higher sampling rate, is required for a smaller propagation distance, which is completely in contradiction to the FFT-ASM. The maximum bandwidth is realized at $x = x_{\max} = M\Delta x$ and $y = 0$ for the FFT-RSI, and at $\alpha = \alpha_{\max} = \pi/\Delta x$ and $\beta = \beta_{\max} = \pi/\Delta y$ for the FFT-ASM. Moreover, if one assumes $\Delta x = \Delta y$ and $k = 2\pi/\lambda$ in air, the requirements for the propagation distances

in the RSI and the ASM can be derived from the above inequalities, which are shown as

$$z_{\text{RSI}} \geq M\Delta x \sqrt{\frac{4\Delta x^2}{\lambda^2} - 1}, \quad (2.91)$$

$$z_{\text{ASM}} \leq M\Delta x \sqrt{\frac{4\Delta x^2}{\lambda^2} - 2}. \quad (2.92)$$

Therefore, for calculation without the problem of aliasing, the above relations in 2.91 and 2.92 have to be fulfilled. They describe the critical propagation distances of numerical simulations with two-dimensional fields propagating in the third dimension. The propagation distance of the FFT-RSI has to be larger than z_{RSI} , while for smaller propagation distances, the sampling interval Δx or the total field size $M\Delta x$ has to be decreased. Conversely, the propagation distance for the FFT-ASM has to be smaller than z_{ASM} , while for larger propagation distances, the sampling interval Δx or the total field size $M\Delta x$ has to be increased.

Furthermore, there are implicit conditions for the spatial sampling interval Δx which can be derived the above inequalities. For the FFT-RSI, its impulse response function will always be sufficiently sampled for $\Delta x \leq \lambda/2$. On the contrary, for the FFT-ASM, the resulting field will always be aliased for $\Delta x \leq \lambda/\sqrt{2}$. Consequently, the FFT-RSI is generally suitable for long propagation distances with a small feature size, while the FFT-ASM is suitable for short propagation distances with a larger feature size.

From a practical perspective, the sampling conditions can be tested with the specifications of the DLAs to be designed. For a focused spot with the shape of an Airy disc, its lateral FWHM is given by the following expression [Wil11]

$$\text{LFWHM} = 0.514 \frac{\lambda}{\text{NA}}. \quad (2.93)$$

For a wavelength of $\lambda = 785$ nm and an NA of 0.7, the lateral FWHM is around 576 nm. However, according to the previous analysis, the FFT-ASM does not allow a sampling interval smaller than $\lambda/\sqrt{2}$ for simulation without aliasing,

which is about 555 nm. Therefore, this creates a barrier for resolving fields with small sizes. Unless the bandwidth of the transfer function is intentionally truncated with the sacrifice of certain high-frequency component, which potentially reduces the accuracy of the simulation. In consequence, either the NA needs to be restricted or the aliasing is involved. Due to the above-mentioned limitations in spatial resolution, the FFT-ASM is not preferable for the high-NA applications in this work, and the FFT-RSI is chosen as the simulation method in Chapter 3.

2.2 Theory of confocal microscopy

In order to predict the performance of the proposed microscope setups, it is important to examine them first in theory. In this section, the theory of scanning confocal microscopy is introduced [Wil84, Kin96]. The principles of its image formation and resolution enhancement are discussed, by which the proposed setup in this work is also analyzed. Specifically, the resolution of a confocal microscope with high-NA illumination and low-NA imaging is evaluated, and the corresponding equations are derived.

2.2.1 Focusing by a thin lens

To start with, consider the simplest situation that a collimated beam of light is focused by a lens as Figure 2.4 shows. The collimated light can be approximated by a plane wave, while the lens can be approximated as an infinitesimal thin element which modulates the field with a transmittance function. Such an approximation is known as the thin lens approximation. For a convex focusing lens, the transmittance function can be written as

$$t_{\text{lens}}(x, y) = P(x, y)e^{-\frac{ik}{2f}(x^2+y^2)}, \quad (2.94)$$

where $P(x, y)$ describes the aperture size. For a circular aperture with a radius of a , it is given as

$$P(x, y) = \begin{cases} 1 & \sqrt{x^2 + y^2} \leq a, \\ 0 & \sqrt{x^2 + y^2} > a. \end{cases} \quad (2.95)$$

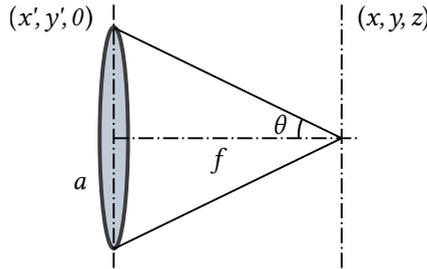


Figure 2.4: A plane wave focused by a thin lens.

Therefore, the field behind the lens under paraxial approximation can be calculated by the Fresnel diffraction integral in Equation 2.51, which is shown as

$$u(x, y, z) = \frac{ke^{ikz} e^{\frac{ik}{2z}(x^2+y^2)}}{i2\pi z} \iint_{-\infty}^{\infty} P(x', y') e^{-\frac{ik}{2f}(x'^2+y'^2)} e^{\frac{ik}{2z}(x'^2+y'^2)} e^{-\frac{ik}{z}(xx'+yy')} dx' dy'. \quad (2.96)$$

The field on the focal plane of the lens can be obtained by substitution of $z = f$ into the above equation, which is simplified as

$$u_f(x, y) = \frac{ke^{ikf} e^{\frac{ik}{2f}(x^2+y^2)}}{i2\pi f} \iint_{-\infty}^{\infty} P(x', y') e^{-\frac{ik}{f}(xx'+yy')} dx' dy'. \quad (2.97)$$

The integral in the above Equation 2.97 is the Fourier transform of the aperture function. For a radially symmetric aperture, the Fourier transform can be

written as the Hankel transform in the following form

$$u_f(r) = \frac{ke^{ikf} e^{\frac{ik}{2f}r^2}}{i2\pi f} \int_0^\infty P(r') J_0\left(\frac{k}{f}rr'\right) 2\pi r' dr', \quad (2.98)$$

where J_0 is the zero-order Bessel function of the first kind, $r' = \sqrt{x'^2 + y'^2}$ and $r = \sqrt{x^2 + y^2}$. For a circular aperture with a radius of a in Equation 2.95, the above integral can be calculated as

$$u_f(v) = -iN_F e^{ikf} e^{i\frac{v^2}{4N_F}} \left[\frac{2J_1(v)}{v} \right], \quad (2.99)$$

where J_1 is the first-order Bessel function of the first kind, N_F is known as the Fresnel number, which is given by

$$N_F = \frac{a^2 k}{2f}, \quad (2.100)$$

and v is the optical coordinate in the radial direction, which is expressed as

$$v = kr \frac{a}{f} \approx kr \sin \theta, \quad (2.101)$$

where $\sin \theta$ denotes the NA in air as shown in Equation 1.1 and Figure 2.4.

By further taking into account the defocused area around the focal plane, the quadratic phase terms in Equation 2.96 no longer cancel with each other and it can be written as

$$u(v,z) = \frac{ke^{ikf} e^{i\frac{v^2}{4N_F}}}{i2\pi f} \int_0^\infty P(r') e^{\frac{-ikr'^2}{2}\left(\frac{1}{f}-\frac{1}{z}\right)} J_0\left(\frac{r'}{a}v\right) 2\pi r' dr'. \quad (2.102)$$

By introduction of the normalized radius and the normalized optical coordinate in the axial direction as

$$\rho = \frac{r'}{a}, \quad (2.103)$$

$$w = ka^2 \left(\frac{1}{f} - \frac{1}{z} \right) \approx \frac{ka^2}{f^2} \delta z \approx 4k \sin(\theta/2) \delta z, \quad (2.104)$$

where $\delta z = z - f$ is the deviation from the focal spot. Equation 2.102 can thus be rewritten as

$$u(v, w) = -iN_{\text{F}} e^{ikz} e^{i\frac{v^2}{4N_{\text{F}}}} \int_0^1 2e^{-\frac{1}{2}iw\rho^2} J_0(v\rho) \rho d\rho. \quad (2.105)$$

It is worthwhile to mention that although the right sides of Equations 2.101 and 2.104 appear to be the approximations of the optical coordinates, the sinusoidal terms are in fact closer to the rigorous solutions under non-paraxial cases as corrections to the original Fresnel diffraction theory [She92]. For the field along the optical axis at $v = 0$, the above integral can be simplified as

$$u(0, w) = -iN_{\text{F}} e^{ikz} e^{-\frac{iw}{4}} \left[\frac{\sin(w/4)}{w/4} \right]. \quad (2.106)$$

And the intensity can be calculated as

$$I(0, w) = |u(0, w)|^2 = N_{\text{F}}^2 \left[\frac{\sin(w/4)}{w/4} \right]^2, \quad (2.107)$$

with $w \propto \delta z$.

2.2.2 PSF of a thin lens

Next, as Figure 2.5 shows, consider that an object with the transmittance of $t(x_1, y_1)$ is located at the first plane. It propagates through a distance d_1 and a thin lens located at the second plane with a focal length of f and a radius

of a . The field right after the lens can be expressed as

$$u_1(\xi_1, \eta_1) = P(\xi_1, \eta_1) e^{-\frac{ik}{2f}(\xi_1^2 + \eta_1^2)} \frac{ke^{ikd_1}}{2\pi i d_1} \iint_{-\infty}^{\infty} t(x_1, y_1) e^{\frac{ik}{2d_1}[(\xi_1 - x_1)^2 + (\eta_1 - y_1)^2]} dx_1 dy_1. \quad (2.108)$$

Then, the field at the image plane (x_2, y_2) can be calculated as

$$\begin{aligned} u_2(x_2, y_2) &= \frac{ke^{ikd_1}}{2\pi i d_2} \iint_{-\infty}^{\infty} u_1(\xi_1, \eta_1) e^{\frac{ik}{2d_2}[(x_2 - \xi_1)^2 + (y_2 - \eta_1)^2]} d\xi_1 d\eta_1 \\ &= -\frac{k^2 e^{ik(d_1 + d_2)}}{4\pi^2 d_1 d_2} \iiint_{-\infty}^{\infty} P(\xi_1, \eta_1) t(x_1, y_1) e^{-\frac{ik}{2f}(\xi_1^2 + \eta_1^2)} \\ &\quad e^{\frac{ik}{2d_1}[(\xi_1 - x_1)^2 + (\eta_1 - y_1)^2]} e^{\frac{ik}{2d_2}[(x_2 - \xi_1)^2 + (y_2 - \eta_1)^2]} dx_1 dy_1 d\xi_1 d\eta_1 \\ &= -\frac{k^2 e^{ik(d_1 + d_2)}}{4\pi^2 d_1 d_2} \iiint_{-\infty}^{\infty} P(\xi_1, \eta_1) t(x_1, y_1) e^{\frac{ik}{2d_1}(x_1^2 + y_1^2)} e^{\frac{ik}{2d_2}(x_2^2 + y_2^2)} \\ &\quad e^{\frac{ik}{2}\left(\frac{1}{d_1} + \frac{1}{d_2} - \frac{1}{f}\right)(\xi_1^2 + \eta_1^2)} e^{-ik\left[\xi_1\left(\frac{x_1}{d_1} + \frac{x_2}{d_2}\right) + \eta_1\left(\frac{y_1}{d_1} + \frac{y_2}{d_2}\right)\right]} dx_1 dy_1 d\xi_1 d\eta_1. \end{aligned} \quad (2.109)$$

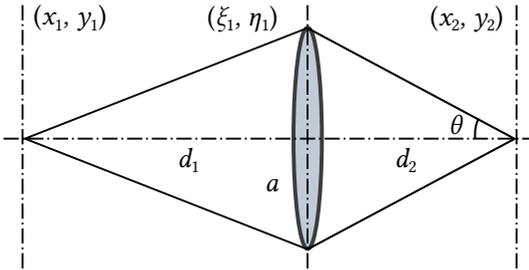


Figure 2.5: Image formation by a thin lens.

When the image plane is in focus, the following lens law is satisfied, which is shown as

$$\frac{1}{d_1} + \frac{1}{d_2} = \frac{1}{f}, \quad (2.110)$$

$$d_2 = M_a d_1, \quad (2.111)$$

where M_a denotes the magnification. Therefore, by substituting the above relations, one can further simplify Equation 2.109 as

$$u_2(x_2, y_2) = - \frac{k^2 e^{ikd_1(1+M_a)} e^{\frac{ik}{2M_a d_1}(x_2^2+y_2^2)}}{4\pi^2 M_a d_1^2} \iint_{-\infty}^{\infty} t(x_1, y_1) e^{\frac{ik}{2d_1}(x_1^2+y_1^2)} h\left(x_1 + \frac{x_2}{M_a}, y_1 + \frac{y_2}{M_a}\right) dx_1 dy_1, \quad (2.112)$$

where the newly introduced function is expressed as

$$h(x, y) = \iint_{-\infty}^{\infty} P(\xi_1, \eta_1) e^{-\frac{ik}{d_1}(\xi_1 x + \eta_1 y)} d\xi_1 d\eta_1. \quad (2.113)$$

The above function is the Fourier transform of the aperture function. For a point source, the object transmittance function can be represented by the Dirac function as follows

$$t(x, y) = \delta(x)\delta(y). \quad (2.114)$$

Consequently, by omitting the prefactor, Equation 2.112 can be directly simplified into

$$u_2(x_2, y_2) \propto h\left(\frac{x_2}{M_a}, \frac{y_2}{M_a}\right). \quad (2.115)$$

Since $h(x_2/M_a, y_2/M_a)$ describes the image of a point, it is denoted as the amplitude point spread function (PSF) or impulse response function. Thereafter,

the same as the derivation from Equation 2.97, the lateral and axial intensity distribution around the focal point has very similar expressions as Equations 2.99 and 2.107, which is expressed by

$$I_2(v, 0) \propto \left[\frac{2J_1(v)}{v} \right]^2, \quad (2.116)$$

$$I_2(0, w) \propto \left[\frac{\sin(w/4)}{w/4} \right]^2. \quad (2.117)$$

The only slight differences lie in the definition of the optical coordinates v and w , which are shown as

$$v = kr_2 \frac{a}{d_2} \approx kr_2 \sin \theta, \quad (2.118)$$

$$w = ka^2 \left(\frac{1}{f} - \frac{1}{d_1} - \frac{1}{d_2 + \delta z} \right) \approx 4k \sin(\theta/2)^2 \delta z, \quad (2.119)$$

where $r_2 = \sqrt{x_2^2 + y_2^2}$ and δz is the axial distance away from the focal point. Moreover, for Equation 2.112, by transformation of the coordinates with $x' = x_1 + x_2/M_a$ and $y' = y_1 + y_2/M_a$, it becomes

$$u_2(x_2, y_2) = - \frac{k^2 e^{ikd_1(1+M_a)} e^{\frac{ik}{2M_a d_1}(x_2^2 + y_2^2)}}{4\pi^2 M_a d_1^2} \iint_{-\infty}^{\infty} t \left(x' - \frac{x_2}{M_a}, y' - \frac{y_2}{M_a} \right) e^{\frac{ik}{2d_1} \left(x'^2 + y'^2 - \frac{2x'x_2}{M_a} - \frac{2y'y_2}{M_a} + \frac{x_2^2}{M_a^2} + \frac{y_2^2}{M_a^2} \right)} h(x', y') dx' dy'. \quad (2.120)$$

With the further assumption that the PSF $h(x', y')$ falls off very quickly for x' and y' away from zero, all the terms in the quadratic phase with x'^2 , y'^2 , $x'x_2$ and $y'y_2$ can thus be eliminated. Subsequently, the above equation can

be approximated by the following form

$$u_2(x_2, y_2) = - \frac{k^2 e^{ikd_1(1+M_a)} e^{\frac{ik}{2M_a d_1} \left(1 + \frac{1}{M_a}\right) (x_2^2 + y_2^2)}}{4\pi^2 M_a d_1^2} \iint_{-\infty}^{\infty} t(x_1, y_1) h\left(x_1 + \frac{x_2}{M_a}, y_1 + \frac{y_2}{M_a}\right) dx_1 dy_1. \quad (2.121)$$

The above integral is indeed a cross-correlation between the object transmittance and the PSF. If the PSF is centrosymmetric, the integral can also be regarded as convolution. And the intensity on the image plane can therefore be calculated as the well-known convolution of the object and the PSF of the imaging system, which is shown as

$$I_2(x'_2, y'_2) = \frac{k^2}{4\pi^2 M_a d_1^2} \left| \iint_{-\infty}^{\infty} t(x_1, y_1) h\left(\frac{x'_2}{M_a} - x_1, \frac{y'_2}{M_a} - y_1\right) dx_1 dy_1 \right|^2, \quad (2.122)$$

with a reversed coordinate $(x'_2, y'_2) = (-x_2, -y_2)$ in the image plane.

2.2.3 PSF of a confocal scanning microscope

With the theories in the above sections, the imaging of a microscope can be analyzed. Two different types of transmitted-light microscopes are shown in Figure 2.6. The conventional microscope in Figure 2.6 (a) has an extended light source. The condenser illuminates the object evenly, and the objective collects the light which passes through the object and forms an image on the image plane.

On the contrary, the confocal scanning microscope in Figure 2.6 (b) has a point light source. The condenser projects a spot on the object and it is imaged by the objective onto the detector. A pinhole is put in front of the detector so that only the intensity at the central point is recorded. The object is scanned step

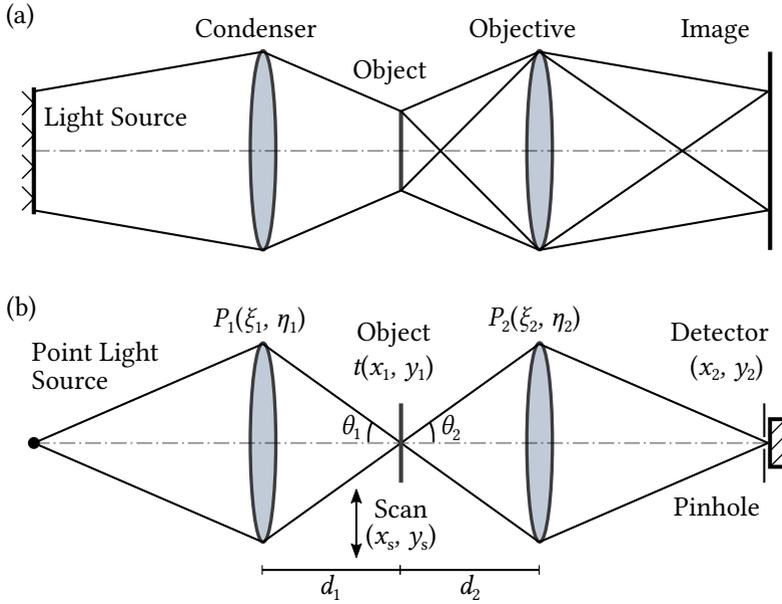


Figure 2.6: Different transmitted-light microscope setups. (a) A conventional microscope. (b) A confocal scanning microscope.

by step and thus the image of the whole object can be reconstructed through the corresponding signal of the detector.

For the conventional microscope in Figure 2.6 (a), its imaging process has no difference than that of a thin lens which has been discussed in Section 2.2.2, since the object is illuminated evenly. The field on the image plane can also be calculated by Equation 2.121. For a point object, the lateral and axial responses of the conventional microscope are also the same as those of a thin lens, which have been shown in Equations 2.116 and 2.117.

For the confocal scanning microscope in Figure 2.6 (b), the field right after the object can be described as the illumination spot of the condenser, which is its PSF, multiplying the object transmittance function

$$u_1(x_1, y_1; x_s, y_s) \propto h_1(x_1, y_1) t(x_1 + x_s, y_1 + y_s), \quad (2.123)$$

where (x_s, y_s) is the scanning position and $h_1(x_1, y_1)$ is the PSF of the condenser. When the defocus is taken into account, the PSF is expressed as

$$h_1(x_1, y_1) = \iint_{-\infty}^{\infty} P_1(\xi_1, \eta_1) e^{-\frac{ik}{d_1}(\xi_1 x_1 + \eta_1 y_1)} e^{-\frac{ik}{2d_1^2}(\xi_1^2 + \eta_1^2)\delta z} d\xi_1 d\eta_1. \quad (2.124)$$

The aperture function $P_1(\xi_1, \eta_1)$ has been expressed in Equation 2.95. Subsequently, the field on the imaging plane (x_2, y_2) can be described as the thin lens imaging formula in Equation 2.121. By omission of the trivial prefactor, the field distribution on the image plane can be written as

$$u_2(x_2, y_2; x_s, y_s) \propto \iint_{-\infty}^{\infty} h_1(x_1, y_1) t(x_1 + x_s, y_1 + y_s) h_2\left(x_1 + \frac{x_2}{M_a}, y_1 + \frac{y_2}{M_a}\right) dx_1 dy_1. \quad (2.125)$$

With a pinhole, only the intensity at $(x_2, y_2) = (0, 0)$ is recorded. Therefore, the intensity response becomes

$$I_2(x_s, y_s) \propto \left| \iint_{-\infty}^{\infty} h_1(x_1, y_1) h_2(x_1, y_1) t(x_1 + x_s, y_1 + y_s) dx_1 dy_1 \right|^2. \quad (2.126)$$

Again, since the PSFs of the objective and the condenser are radially symmetric, the above integral can be regarded as a convolution as discussed in the previous section. For a point object with $t(x_1 + x_s, y_1 + y_s) = \delta(x_1 + x_s, y_1 + y_s)$, the above integral can be easily simplified as

$$I_{\text{Point}}(x_s, y_s) \propto |h_1(x_s, y_s) h_2(x_s, y_s)|^2, \quad (2.127)$$

which indicates that the overall PSF for the confocal system is the product of $h_1(x_s, y_s)$ and $h_2(x_s, y_s)$. Similar to the analysis in Section 2.2.1, by use of the normalized optical coordinates, the following expressions for the radial and axial responses can be derived as

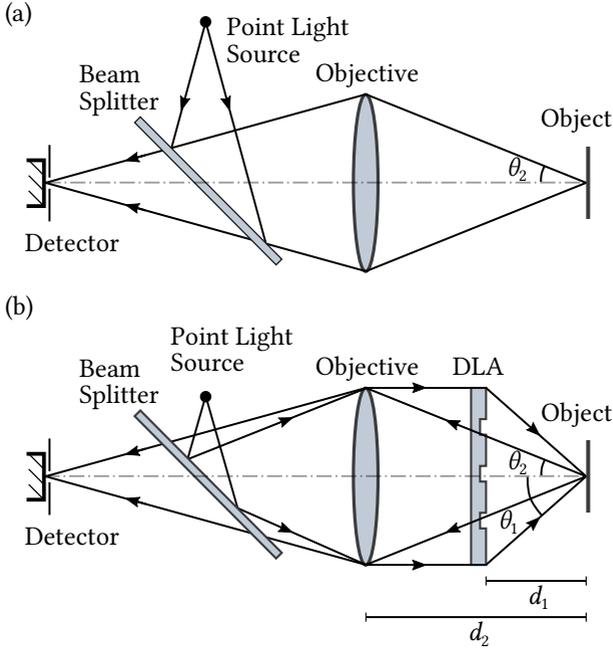


Figure 2.7: Reflected-light microscope setups. (a) Traditional confocal microscope. (b) DLA-based confocal microscope.

$$I_{\text{Point}}(v, 0) \propto \left[\frac{2J_1(v_1)}{v_1} \right]^2 \left[\frac{2J_1(v_2)}{v_2} \right]^2, \quad (2.128)$$

$$I_{\text{Point}}(0, w) \propto \left[\frac{\sin(w_1/4)}{w_1/4} \right]^2 \left[\frac{\sin(w_2/4)}{w_2/4} \right]^2, \quad (2.129)$$

where the optical coordinates are expressed as

$$v_{1,2} = k \sqrt{x_s^2 + y_s^2} \sin(\theta_{1,2}), \quad (2.130)$$

$$w_{1,2} = 4k \sin(\theta_{1,2}/2)^2 \delta z. \quad (2.131)$$

Next, consider a reflection confocal microscope in Figure 2.7 (a). Its image formation is exactly the same as the transmission confocal microscope. Therefore, the intensity response of such a setup also satisfies Equation 2.126, while h_1 and h_2 in the expression are identical in this case. On the other hand, the DLA-based confocal microscope in Figure 2.7 (b) utilizes the DLA as the condenser. The DLA can have a higher NA than the objective to increase the overall resolution according to Equations 2.128 and 2.129 when a point object is imaged.

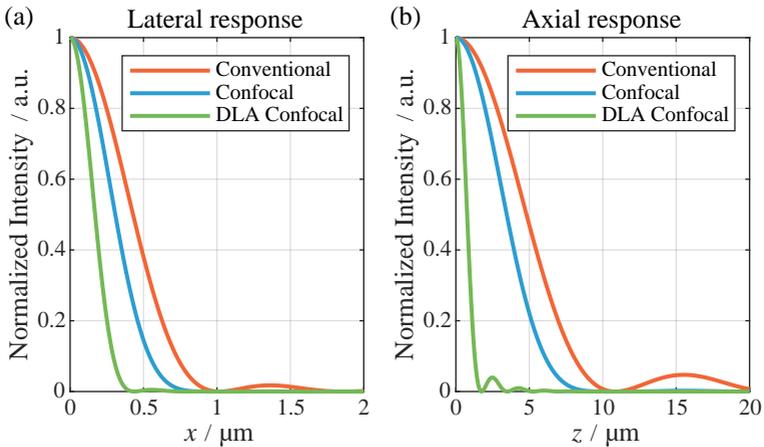


Figure 2.8: Comparison of a conventional and a confocal microscope with a 0.3 NA objective, and a DLA-based confocal microscope with a 0.3 NA objective and a 0.7 NA DLA when a point object is imaged. (a) Lateral responses. (b) Axial responses.

As an example, Figure 2.8 shows the lateral and axial response of a point object imaged by the conventional microscope, the confocal microscope, and the DLA-based confocal microscope. The NA of the objective and the DLA is set to 0.3 and 0.7 respectively. The wavelength is $0.5 \mu\text{m}$. The result clearly shows that with the same objective, the DLA-based confocal microscope provides a better resolution than the conventional microscope and the traditional confocal microscope in both lateral and axial directions when a point object is imaged. However, for a plane object, the results are different. For a perfect plane

reflector, the object function becomes

$$t(x, y) = 1. \quad (2.132)$$

Now consider only the axial response along the optical axis with $x_2 = y_2 = 0$ for the conventional microscope, the intensity response is expressed by

$$\begin{aligned} I_{\text{Plane}}(0, w) &\propto \left| \iint_{-\infty}^{\infty} h_2(x_1, y_1) dx_1 dy_1 \right|^2 \\ &\propto \left| \iint_{-\infty}^{\infty} \iint_{-\infty}^{\infty} P_2(\xi_2, \eta_2) e^{-\frac{ik}{d_2}(\xi_2 x_1 + \eta_2 y_1)} e^{-\frac{ik}{2d_2^2}(\xi_2^2 + \eta_2^2)\delta z} d\xi_2 d\eta_2 dx_1 dy_1 \right|^2 \\ &\propto \left| \iint_{-\infty}^{\infty} P_2(\xi_2, \eta_2) e^{-\frac{ik}{2d_2^2}(\xi_2^2 + \eta_2^2)\delta z} \delta(\xi_2) \delta(\eta_2) d\xi_2 d\eta_2 \right|^2 \\ &\propto 1, \end{aligned} \quad (2.133)$$

which means that the conventional microscope cannot distinguish the height variation of a plane object at all. Similarly, for the confocal microscope, the axial intensity response with $x_s = y_s = 0$ is given by

$$\begin{aligned} I_{\text{Plane}}(0, w) &\propto \left| \iint_{-\infty}^{\infty} h_1(x_1, y_1) h_2(x_1, y_1) dx_1 dy_1 \right|^2 \\ &\propto \left| \iiint_{-\infty}^{\infty} P_1(\xi_1, \eta_1) P_2(\xi_2, \eta_2) e^{-\frac{ik}{2d_1^2}(\xi_1^2 + \eta_1^2)\delta z} e^{-\frac{ik}{2d_2^2}(\xi_2^2 + \eta_2^2)\delta z} \right. \\ &\quad \left. \delta\left(\frac{\xi_1}{d_1} + \frac{\xi_2}{d_2}\right) \delta\left(\frac{\eta_1}{d_1} + \frac{\eta_2}{d_2}\right) d\xi_1 d\eta_1 d\xi_2 d\eta_2 \right|^2 \\ &\propto \left| \iint_{-\infty}^{\infty} P_1\left(\frac{d_1}{d_2}\xi_2, \frac{d_1}{d_2}\eta_2\right) P_2(\xi_2, \eta_2) e^{-\frac{ik}{d_2^2}(\xi_2^2 + \eta_2^2)\delta z} d\xi_2 d\eta_2 \right|^2. \end{aligned} \quad (2.134)$$

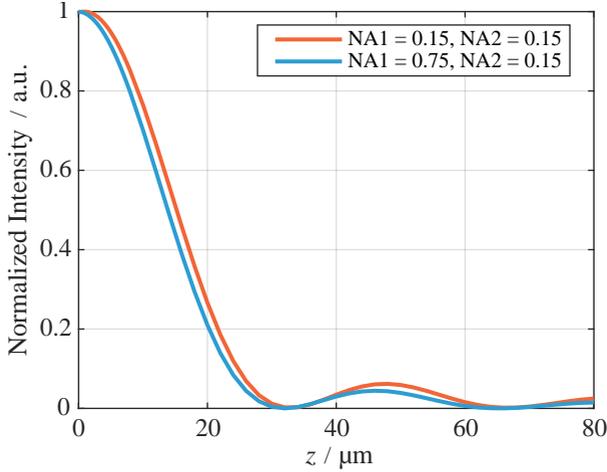


Figure 2.9: Axial intensity responses of confocal systems for a plane object by illumination and imaging with different NAs. With an illumination NA of 0.75 and an imaging NA of 0.15, the HWHM of the axial response is 13.85 μm . With both an illumination NA and an imaging NA of 0.15, the HWHM of the axial response is 15.22 μm .

Without loss of generality, suppose that the illumination has a larger NA than the objective, which is given by $a_1/d_1 > a_2/d_2$. Consequently, the first aperture function P_1 in the above integral can be neglected. Equation 2.134 can thus be calculated in the same way as Equation 2.102 by introducing the normalized optical coordinates. The result is given by

$$I_{\text{Plane}}(0, w) \propto \left[\frac{\sin(w_2/2)}{w_2/2} \right]^2, \quad (2.135)$$

where w_2 is defined by Equation 2.131. From the above analysis, it can be seen that unlike the conventional microscope, the confocal microscope has the depth discerning capability and thus it can be used for surface topography measurement. However, if the illumination and the imaging have different NAs, the response for a plane object is only determined by the one with the lower NA. That is to say, for the DLA-based confocal microscope, the high-NA illumination by the DLA has no contribution to the axial measurement sensitivity when a plane is measured.

It should be noted that all the previous analysis in this section is based on the Fresnel diffraction formula for the ease of analytical calculation, which is less accurate in high-NA situations. However, it does provide a correct qualitative understanding for the responses of the microscope systems. For example, under a more rigorous scalar diffraction simulation with FFT-RSI for the measurement of a plane, the high-NA illumination of the DLA-based confocal microscope still helps to gain marginally better resolution than the traditional confocal microscope [Li19a]. As shown in Figure 2.9, with an illumination NA of 0.75 and an imaging NA of 0.15, the half width at half maximum (HWHM) of the plane response is $13.85\ \mu\text{m}$, while the HWHM is $15.22\ \mu\text{m}$ with both the illumination and the imaging NA of 0.15. Nevertheless, such a gain is not substantial and the width of the intensity response is still dominated by the low-NA objective, which is predicted by the analytical formula expressed in Equation 2.135.

In the next chapter, the detailed design and simulation of the DLAs and the systems are introduced. New concepts are also proposed to increase the axial measurement sensitivity and to solve the above-mentioned deficiency in surface measurement for the See-through DLAs.

3 Design and Simulation

In this chapter, in order to solve the restriction between the FOV and the resolution in traditional confocal microscopes, two DLA concepts are proposed. Unlike the DLAs in previous research, they are not limited solely in transmitted-light microscopes anymore, and they can be successfully implemented in reflected-light configurations to increase the overall SBPs of the measurement systems.

In Section 3.1, the See-through DLAs are proposed. They aim to reduce the disturbance when light passes through them by increase of the zero-order diffraction efficiency. They can be used with low-NA objectives while still maintaining high resolution when point objects are measured.

Furthermore, in Section 3.2, the Direct-imaging DLAs are proposed. They aim to act as high-NA lens arrays by not only producing the illumination spots, but also imaging the spots back to the image planes by themselves. In consequence, they can also measure surface topography with high resolution.

In the following sections, the two DLA concepts will be introduced in detail. Their advantages and shortcomings will be discussed. The simulation and optimization processes of the DLAs will be thoroughly explained.

3.1 See-through DLA design

As is briefly discussed in Section 1.2, the DLAs in previous research have only been used as pure illuminators in transmitted-light microscope configurations as shown in Figure 1.2 (a), which are substantially not preferable for applications like surface metrology and fluorescence microscopy [Plo99, Rus09].

In order to employ them in reflected-light microscopes such as the one in Figure 1.2 (b), the produced illumination spots have to be imaged through the DLAs again. There is generally no workaround to avoid the transmission through the DLAs in the imaging path. Using a beam splitter to reflect the illumination spots from the side of the objective is not possible, as long as the focusing light cones have the half angles over 45 degrees, let alone the short working distances and the large measurement areas. Therefore, a natural way to overcome this problem is to reduce the disturbance caused by the DLA in the imaging path, or in other words, to make the DLA more transparent which does not obstruct the imaging of the objective. Such a DLA concept is called the See-through DLA in this work and it will be explained in more details in the following contents.

3.1.1 System overview

The schematic of the confocal microscope based on the See-through DLA is shown in Figure 3.1. A point light source is reflected by a beam splitter and collimated by an objective. The DLA is illuminated by the collimated light and it produces an array of focused spots. The spots are imaged by the objective through the DLA onto the camera sensor.

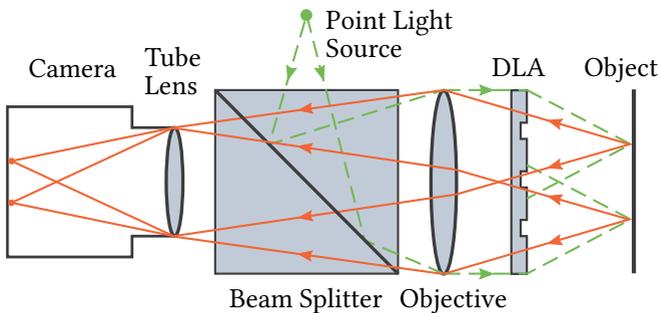


Figure 3.1: Schematic of the confocal microscope setup based on the See-through DLA.

The DLA is a kind of DOE which modulates the light passing through it and forms certain patterns after a working distance. The DOE is realized either by a phase mask or by an amplitude mask. A phase mask is usually a piece of etched glass with micro-steps of different heights at each position. Light passing through these micro-steps experiences different optical path lengths. In such a way, the phase of light can be modulated by the different thickness of glass at each position. In this way, one can control the phase distributions and realize the desired projection patterns according to the design of the mask by diffraction simulation.

On the other hand, an amplitude mask is usually a piece of glass with reflective coatings, e.g., chromium, at selected positions. Hence, the amplitude of light can be controlled by allowing light transmission only at certain positions. The phase and amplitude masks can also be used in reflection mode for modulation of the reflected light. The most common type of DOE is the transmission binary phase mask, because amplitude masks can only utilize a small part of the light due to the reflective coating. Besides, binary concave and convex structures are much easier to produce than multi-level steps by lithography. Therefore, the DLAs in this work are specifically realized by binary phase masks in both the design processes and the experiments.

To use the DLAs for fast scanning with high resolution, they have to produce dense spot arrays with high NA. This is achieved by the concept of overlapping apertures. As shown in Figure 3.2, the spot array is produced by the DLA which is composed of periodic unit cells. The initial target array is constructed and each spot creates a spherical-like wave which propagates through a certain working distance to the DLA plane. However, on the DLA plane, the waves created by each spot are no longer restricted in the unit cells. They are overlapped with each other and equivalently form a lens array with overlapping apertures. The phase of the field is binarized which results in a binary pattern for making a binary phase mask. Hence, when parallel light passes through the phase mask, it projects a spot array similar to the initial target array after the working distance. It is worth noting that the projected field differs from the target field due to the binarization, and such difference can be reduced by optimization of the phase mask pattern. Nevertheless, in this

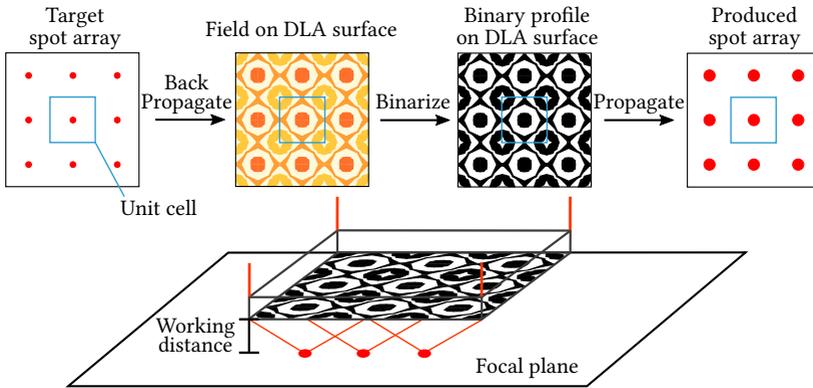


Figure 3.2: The DLA with overlapping apertures.

way, a single spot in the produced array not only receives light from the unit cell right above it, but it also receives contributions from all unit cells around the spot. In other words, the aperture of the lens which produces the spot is not limited in the small unit cell. It is indeed overlapped with the adjacent ones. In consequence, the NA of the spot is no longer restricted by the size of the unit cell or the pitch of the array anymore. A dense spot array with high NA can thus be generated in this way.

Previous works also have similar ideas for the overlapping apertures, but with different design approaches. For example, Wu et al. [Wu10] use a pinhole array to generate the initial target spot array with a collimated laser beam and physically record the DLA field by the volumetric interference in a holographic plate, which is similar to a thick sheet of analog film. Hulsken et al. [Hul12] and Stenau et al. [Ste16] use the FFT-ASM to simulate the diffraction propagation and to design the phase masks with sub-wavelength features. They abandon the zero-padding and intentionally use the periodicity and overlapping as explained in Section 2.1.3.3, so that the simulation of a periodic spot array can be replaced by the simulation of only one unit cell, which reduces the computation cost. However, in order to design the DLAs in a reflected-light configuration in Figure 3.1, simulation of the whole imaging process is necessary to increase the SNR in the image. For this reason, none

of the above-mentioned approaches are suitable. The volumetric holography approach is difficult to control. The FFT-ASM without zero-padding can only simulate periodic structures and it is not capable of simulating the imaging by a thin lens. Therefore, a new simulation framework based on the FFT-RSI needs to be established for the design of the See-through DLAs.

3.1.2 Design process

To make the See-through DLAs suitable for the reflected-light microscopes, the idea is to make them more transparent or namely to increase their zero-order diffraction efficiency. Traditionally, this can be done by iterative pixel-wise optimization after generation of the initial pattern of the phase mask. That is to say, all the pixels of the phase pattern are treated as variables. The value of each pixel is varied to check if the result is closer to the target. However, such a method is not suitable for the DLAs in this work, since the large area and the small feature size lead to a large number of pixels. For example, for a DLA with an area of $2\text{ mm} \times 2\text{ mm}$ and a feature size of $0.4\text{ }\mu\text{m}$, there are 2.5×10^7 pixels in total, which result in way too many variables for the optimization.

Therefore, a physically based approach has been used instead to increase the zero-order diffraction efficiency, which is done by adding a plane-wave component to the original field distribution of the DLA. Such a method is inspired by the multi-functional DOEs [Vij15], whose key idea is the superposition of different field distributions. By simply adding different fields generated from different target light distributions, all the target patterns can be generated with

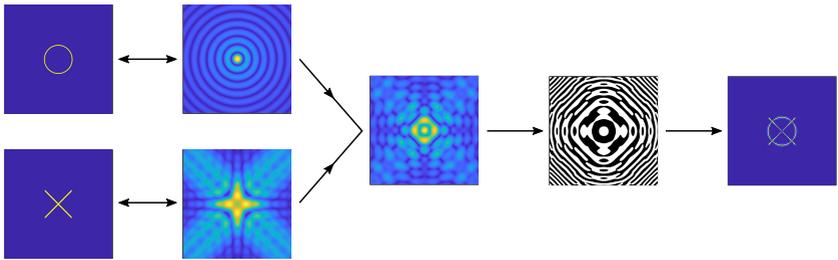


Figure 3.3: Superposition of fields for multi-functional DOEs.

one single DOE. For example, as the simulation in Figure 3.3 shows, if one field can generate a circle and another one can generate a cross after propagation through a certain distance, by simply adding the two fields together, the resulting binary phase pattern can generate a circle and a cross at the same time. In such a way, a single piece of DOE can have more flexible functionalities.

Similarly, the See-through DLA is made by combination of the original DLA with a plane wave, which effectively increase the zero-order diffraction efficiency of the original DLA. The detailed design process can be found in Figure 3.4. First, an ideal target spot array u_{array} is constructed by replication of the spots in a grid. The spots are represented by Airy discs to limit the bandwidth of the field. The resulting field can be expressed by the following equation

$$u_{\text{array}}(x, y) = \sum_{m,n} \frac{2J_1(k_0\text{NA}\sqrt{(x-mT)^2+(y-nT)^2})}{k_0\text{NA}\sqrt{(x-mT)^2+(y-nT)^2}}, \quad (3.1)$$

where NA is the designed numerical aperture, T is the designed pitch of the grid, m, n are integers which enumerate of the spots in the array, and $k_0 = 2\pi/\lambda$ is the wave number in air.

Second, the field distribution u_p of the DLA is obtained by the propagation of u_{array} after a working distance d_0 by the Rayleigh-Sommerfeld integral described in Equation 2.34, which is expressed as

$$u_p = S\{u_{\text{array}}, k_0, d_0\}, \quad (3.2)$$

where the operator $S\{\cdot\}$ denotes the propagation by the Rayleigh-Sommerfeld integral. In practice, the field is discretized into a matrix with a defined spatial interval according to Equation 2.63, and the FFT-RSI method in Equation 2.71 is used for the diffraction simulation. The following simulations are all carried out in this way, and it will not be further mentioned for simplicity. Afterwards, a plane-wave component is added to u_p , which is the key step to increase the zero-order diffraction of the DLA, and it is expressed as

$$u_S = u_p + W_S, \quad (3.3)$$

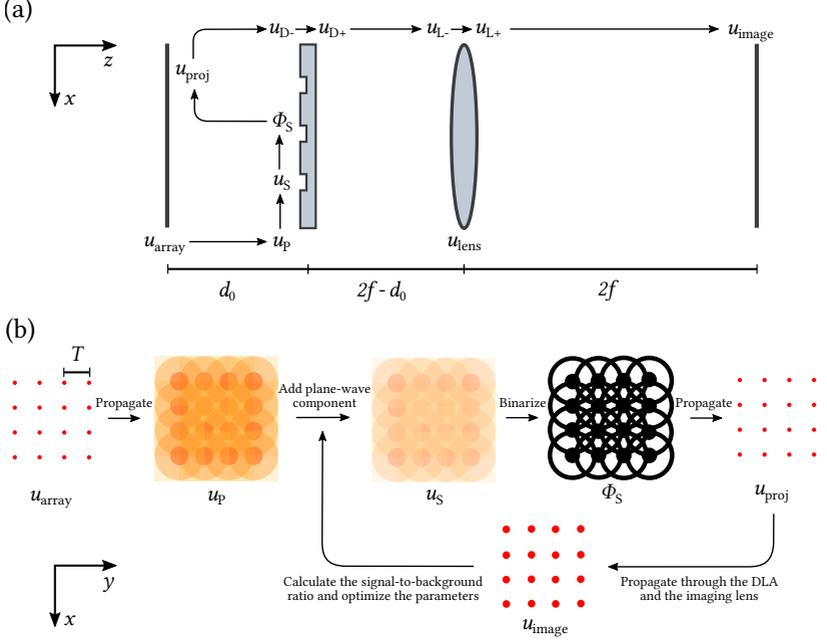


Figure 3.4: Design process of the See-through DLAs. (a) Pipeline for the image simulation. (b) Flow chart for the design and optimization.

where W_S is a constant which represents the amount of the plane-wave component. Then, the phase of the field distribution u_s is extracted and binarized by the following equation

$$\phi_S = \text{mod} \left(\left\lfloor \frac{\arg(u_s)}{\pi} + B \right\rfloor, 2 \right) \pi, \quad (3.4)$$

where $\arg(\cdot)$ denotes extraction of the phase angle from the complex field, $\lfloor \cdot \rfloor$ represents the floor operator, and B is a binarization factor between 0 and 1. The phase distribution $-\phi_S$ also represents the binary pattern which will be printed on the phase mask for production. The minus sign is necessary to reverse the propagation direction, although it does not matter for a binary phase pattern with only 0 and π .

Next, the whole imaging process is simulated. To start with, the binary DLA pattern was projected to the object plane, which is expressed as

$$u_{\text{proj}} = S\{e^{-i\phi_s}, k_0, d_0\}, \quad (3.5)$$

where u_{proj} is the simulated spot array produced by the binary phase mask of the DLA, and $e^{-i\phi_s}$ is field distribution when light passes through the phase mask. Suppose that the spot array illuminates a mirror surface and it is reflected back to the DLA plane again. Similarly, the field right before the DLA plane is written as

$$u_{\text{D-}} = S\{u_{\text{proj}}, k_0, d_0\}. \quad (3.6)$$

Afterwards, the field again passes through the binary DLA pattern and the phase mask is also assumed to be a thin element with negligible thickness. It should be noted that it is possible to include the glass thickness of the phase mask in the simulation by changing the wave number and the propagation distance. However, compensation for the glass thickness is more related to the properties of the objective. Such an additional step will introduce aberrations in the imaging path and make the focal plane of the lens more difficult to define in the simulation. Hence, the glass thickness is neglected for simplicity reason. The field right after the phase mask can thus be calculated as

$$u_{\text{D+}} = u_{\text{D-}} \cdot e^{-i\phi_s}. \quad (3.7)$$

Afterwards, assume that the spots are imaged by a thin lens in a $4f$ configuration with a one-to-one magnification, where f is the focal length of the thin lens. Consequently, in order to keep the spots in focus, the total optical path length needs to be $2f$ and thus the distance between the DLA surface and the thin lens becomes $2f - d_0$. Then, the field right before the thin lens is represented by

$$u_{\text{L-}} = S\{u_{\text{D+}}, k_0, 2f - d_0\}. \quad (3.8)$$

The lens is under the thin-element approximation and its thickness is neglected. Subsequently, the field right after the lens is simply a multiplication with its transmittance function in Equation 2.94, which is expressed as

$$u_{L+} = u_{L-} \cdot P(x, y)e^{-\frac{ik_0}{2f}(x^2+y^2)}, \quad (3.9)$$

where $P(x, y)$ describes the aperture size of the lens which has been expressed in Equation 2.95. By adjustment of the aperture and the focal length, the NA of the thin lens can be defined as required. Finally, the field on the image plane is obtained by

$$u_{\text{image}} = S\{u_{L+}, k_0, 2f\}. \quad (3.10)$$

3.1.3 Simulation and optimization

The whole imaging process of the See-through DLA has been presented in the above subsection. Based on that, the parameters in the above equations can be optimized. First, a peak-to-background ratio (PBR) is defined, which is shown by the following equation

$$\text{PBR}(m, n) = \frac{I_{\text{peak}}(m, n)}{I_{\text{mean}}(m, n)}, \quad (3.11)$$

where I_{peak} is the intensity of the central pixel of the spot in the (m, n) unit cell on the image plane, and I_{mean} is the average background intensity of the unit cell excluding the circular area of the central spot with a diameter of 3 times its FWHM. The PBR represents the SNR in the simulated image, and the target of the optimization is to maximize it. For each unit cell, its PBR can be calculated accordingly. However, due to the inherit off-axis aberration of a singlet lens, only the central part of the spot array can be clearly imaged. In practice, the size of the low-aberration area is related to the defined NA of the lens. Thus, only the PBRs of the central $N' \times N'$ unit cells are calculated. Their

average value is defined as the optimization merit, which is expressed as

$$\text{PBR}_{N'} = \sum_{m,n} \frac{\text{PBR}(m, n)}{N'^2}, \quad (3.12)$$

where N' is chosen empirically according to the simulated image. The iterative optimization process is briefly shown in Figure 3.4 (b). The parameters are updated in every loop in order to maximize the above mentioned merit. On the one hand, some parameters are set to be fixed according to the specifications of the real system, including the wave numbers, the number of the spots, the pitch of the array T , the focal length of the imaging objective f and the aperture of the objective. On the other hand, the other parameters are left to be variables in the optimization, including the working distance d_0 , the weight of the plane-wave component W_S and the binarization factor B . Subsequently, the whole optimization argument is expressed as

$$(B, d_0, W_S)_{\text{opt}} = \arg \max_{B, d_0, W_S} \text{PBR}_{N'}(B, d_0, W_S), \quad (3.13)$$

where $(B, d_0, W_S)_{\text{opt}}$ denotes the optimum parameters and $\arg \max$ means argument of the maximum. Among the three variables, only the global optimum value of W_S can be easily found by the gradient descent method when the other two are fixed. This can be well explained that a small amount of plane-wave component leads to larger disturbance, i.e., higher noises, while a large amount of plane-wave component leads to fainter spots, i.e., weaker signals. So, there should be a balanced point to achieve the highest SNR. On the contrary, for B and d_0 , they are highly coupled with no explicit relations and the PBR is rather a none-convex function for the two variables with plenty of local maximums, which is common in optical design. Therefore, the strategy for the optimization is to treat them separately.

As Figure 3.5 shows, first, initial values are chosen for the parameters, for example, $d_0 = 1000 \mu\text{m}$, $B = 0.5$ and $W_S = 0$. Second, the plane-wave component weight W_S is first optimized with the gradient descent method to get a best value for the initial d_0 and B . Third, with the updated W_S , a

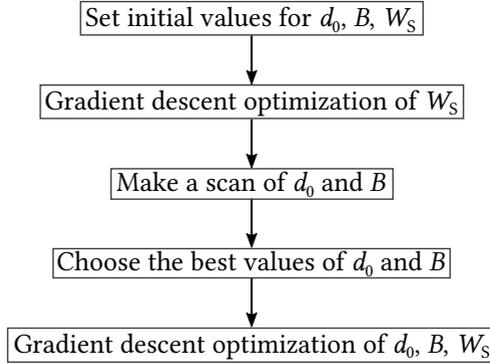


Figure 3.5: Optimization strategy for the See-through DLAs.

scan of d_0 and B is made within a defined region, for example, d_0 is chosen from $900\ \mu\text{m}$ to $1100\ \mu\text{m}$ with an interval of $5\ \mu\text{m}$, and B is chosen from 0 to 0.9 with an interval of 0.1. Then, d_0 and B is set to the values with the maximum average PBR in the scan. At last, a gradient descent optimization is again performed from the new starting point for all the parameters to find the optimum values of them in the defined region.

As an example, a See-through DLA is optimized to generate a 21×21 spot array with a pitch of $20\ \mu\text{m}$. The simulation wavelength is $785\ \text{nm}$, the feature size is $0.2\ \mu\text{m}$ and the size of the total computation window is $2.4\ \text{mm}$. The NA of the singlet imaging lens is 0.3 with a focal length of $1.85\ \text{mm}$. The designed NA for the spot array projected by the DLA is 0.7. The other variables are set according to the above-mentioned optimization process with a binarization factor B of 0.71, a working distance d_0 of $1091\ \mu\text{m}$ and a plane-wave component weight W_S of 20.7.

The simulation results are shown in Figure 3.6. In Figure 3.6 (a) and (b), the DLA is able to project a spot array as designed. The FWHM of the central spot projected on the object plane is $0.594\ \mu\text{m}$, which corresponds to an NA of 0.68 according to Equation 2.93. Moreover, the spots imaged by the objective on the image plane are shown in Figure 3.6 (c) and (d). It is shown that through the imaging process in Figure 3.4 (a), the spots are still clearly visible when

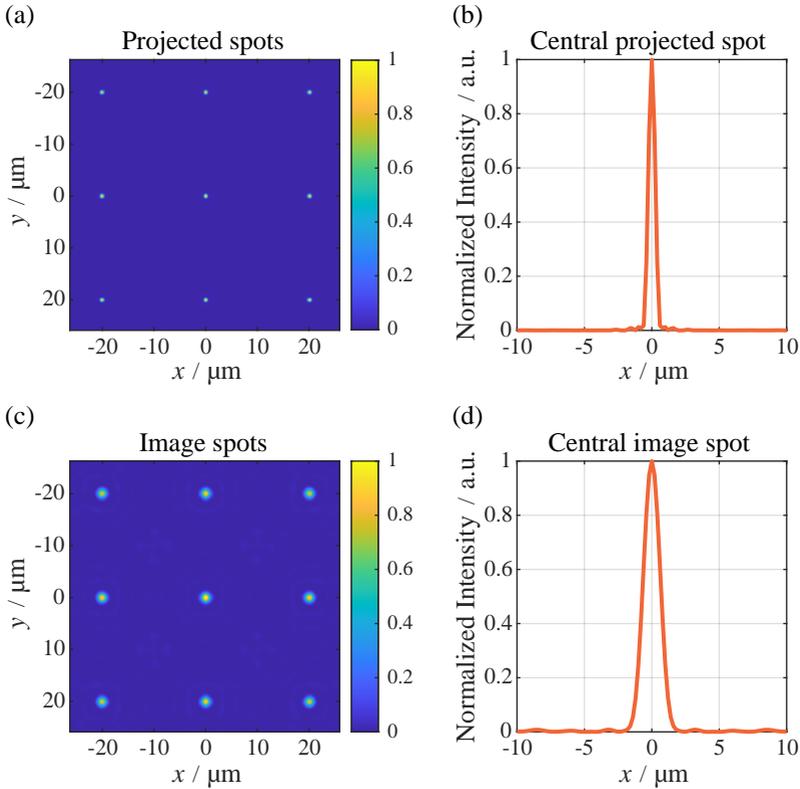


Figure 3.6: Simulation of spots generated and imaged through a See-through DLA. (a) The central 3×3 grid in the 21×21 spot array produced by the DLA on the object plane. (b) Cross section of the central spot on the object plane at $y = 0$. The FWHM is $0.594 \mu\text{m}$. (c) The central 3×3 grid in the 21×21 spot array through the objective on the image plane. (d) Cross section of the central spot on the image plane at $y = 0$.

they are imaged through the designed See-through DLA without too much disturbance. The average PBR of the central 5×5 unit cells is 236. It is worth noting there are slight deviations of the brightness of the spots at different positions due to different extents of overlapping of the apertures, which can be compensated by adjusting the initial intensity of the spots in the ideal spot array u_{array} if necessary.

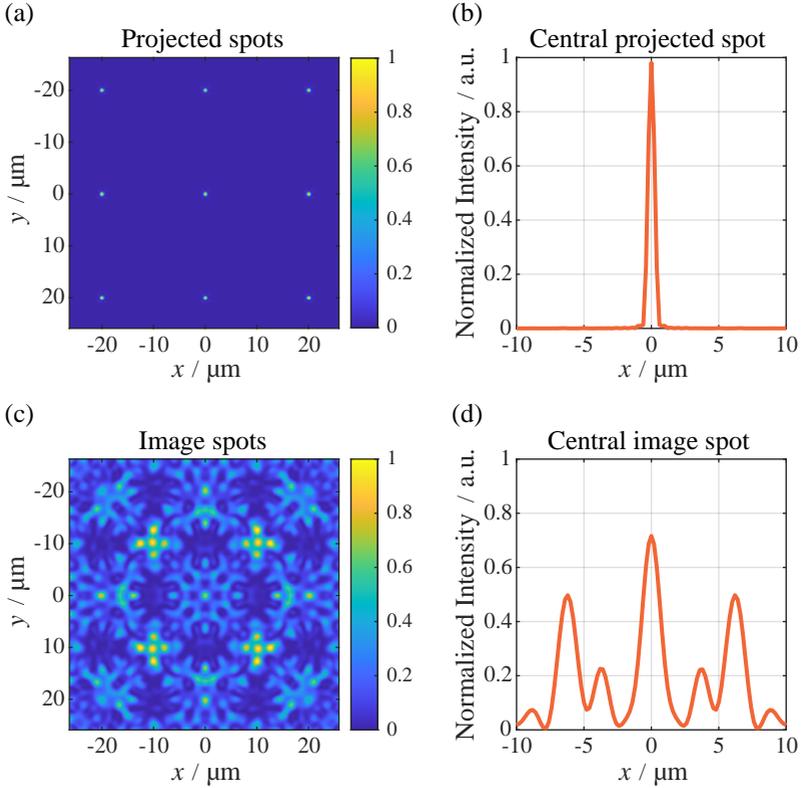


Figure 3.7: Simulation of spots generated and imaged through a DLA without the plane-wave component. (a) The central 3×3 grid in the 21×21 spot array produced by the DLA on the object plane. (b) Cross section of the central spot on the object plane at $y = 0$. The FWHM is $0.579 \mu\text{m}$. (c) The central 3×3 grid in the 21×21 spot array through the objective on the image plane. (d) Cross section of the central spot on the image plane at $y = 0$.

By contrast, Figure 3.7 shows the simulation results for an original DLA without any plane-wave component. All the parameters are the same as above except that W_S is set to zero. The FWHM of the central spot projected on the object plane in Figure 3.7 (b) is $0.579 \mu\text{m}$, which corresponds to an NA of 0.697 and is very close to the designed value. However, by the simulated image in Figure 3.7 (c), the spots on the image plane are no longer visible. They are totally covered

by noise which originates from the disturbance of the DLA. By comparing the simulation results in Figure 3.6 and 3.7, it is found that the introduction of the plane-wave component leads to a slight increase of the projected spot size, but also a significant increase of the SNR of the image when the spots are imaged through the DLA.

In summary, the proposed method has been shown to successfully increase the zero-order diffraction efficiency of the phase mask in the simulation. Accordingly, the See-through DLAs can be well adapted in the reflected-light microscope configuration as Figure 3.4 shows. Such a setup can improve the resolution and increase the overall SBP of the system by the high-NA spots produced by the DLAs. In consequence, high-resolution large-area measurement of opaque surfaces or fluorescent samples are made possible by the See-through DLAs.

3.2 Direct-imaging DLA design

In the previous section, it is shown that the proposed See-through DLAs can be applied in reflected-light microscopes and increase the lateral resolution of the objective by confocal scanning. However, as has been already discussed in Section 2.2.3, when the object is a plane or an opaque surface, the axial measurement sensitivity of such a setup in Figure 3.4 is still fundamentally limited by the low-NA objective.

In order to increase the axial sensitivity for 3D surface measurement, the Direct-imaging DLAs are proposed in this section. They act exactly as finite-conjugate high-NA lens arrays, and they can provide high resolution for 3D confocal surface measurement in both the lateral and the axial directions. They are designed by the superposition of the field distributions of different lenses with overlapping in the arrays. The detailed design and optimization processes are presented in the following sections.

3.2.1 System overview

The schematic of the confocal microscope based on the Direct-imaging DLA can be found in Figure 3.8. Similarly as in the See-through setup, a point light source is reflected by a beam splitter. Light is collimated by an objective, and the collimated plane wave illuminates the DLA. Then, the DLA projects an array of focused spots on the object plane. This time, the produced spots are not imaged by the objective, but they are directly imaged by the DLA itself to an intermediate image. Afterwards, the intermediate image is collected by the objective and the tube lens to the camera sensor.

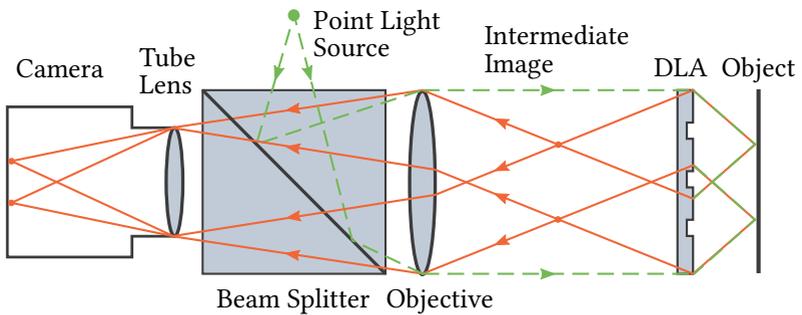


Figure 3.8: Schematic of the confocal microscope setup based on the Direct-imaging DLA.

In such a setup, the Direct-imaging DLA acts exactly as an array of finite-conjugate high-NA objectives. The imaging of the object is solely done by the DLA itself and the objective is used to collect the intermediate image. As a result, when a plane object is measured, the axial measurement sensitivity is only determined by the DLA, which is designed to have a much larger NA than the objective. In this way, the setup can perform high-NA multi-spot confocal measurement with both point-like objects and plane surfaces.

According to Figure 3.8, the Direct-imaging DLAs need to realize two functions. One is to project an array of spots with plane-wave illumination, and the other one is to image individual spots onto the intermediate image plane. This is again realized by the multi-functional DOEs. However, the field components

are different than those of the See-through DLAs in the previous section, which will be introduced in the following contents.

3.2.2 Design process

The design of the Direct-imaging DLAs begins with construction of the unit element. As shown in Figure 3.9, the unit element is composed of the field distributions of two different diffractive lenses. One is the illumination lens which has the focal length of d_0 , and it focuses plane-wave illumination to a spot on the object plane. The other one is the imaging lens which has the focal length of $d_0 d'_0 / (d_0 + d'_0)$ and it images the produced spot back to an intermediate image. Similarly, by the principle of the multi-functional DOEs, the two lenses can function at the same time by simply adding their field distributions together.

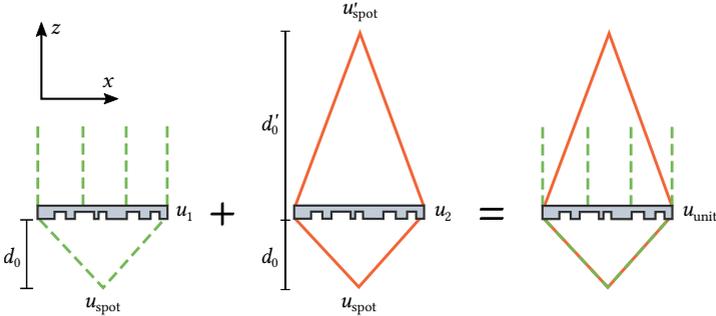


Figure 3.9: Construction of the unit element by the superposition of two field distributions.

The field distribution of the illumination lens u_1 is formed by the propagation of the designed focused spot u_{spot} through a certain working distance d_0 . As previously mentioned, u_{spot} is also set to an Airy disc to limit the spatial bandwidth with the expression of

$$u_{\text{spot}}(x, y) = \frac{2J_1(k_0 \text{NA} \sqrt{x^2 + y^2})}{k_0 \text{NA} \sqrt{x^2 + y^2}}. \quad (3.14)$$

Subsequently, u_1 can be calculated by the RSI as

$$u_1 = S\{u_{\text{spot}}, k_0, d_0\}. \quad (3.15)$$

On the other hand, the field distribution u_2 for the imaging lens is calculated by the propagation of two spots from two sides, which is given by

$$u_2 = \frac{S\{u'_{\text{spot}}, k_0, d'_0\}^*}{S\{u_{\text{spot}}, k_0, d_0\}}, \quad (3.16)$$

where $(\cdot)^*$ denotes the complex conjugate, u'_{spot} is simply another Airy disc as expressed Equation 3.14 with a different numerical aperture of NA' , and d'_0 is the designed distance from the DLA plane to the intermediate image shown in Figure 3.9. The complex conjugate is necessary to reverse the propagation direction of the divergent wave.

Next, the field distribution of the unit element can be obtained by the superposition of the two lenses u_1 and u_2 . The process is similar to the superposition of the plane-wave component in Equation 3.3, which is expressed as

$$u_{\text{unit}} = u_1 + W_{\text{D}} u_2, \quad (3.17)$$

where W_{D} is the ratio between the illumination lens and the imaging lens. By the above-mentioned process, the calculated unit element can produce a focused spot by the plane-wave illumination, and then it can image the spot back to the intermediate plane by itself.

Afterwards, the pattern of the DLA is calculated as Figure 3.10 shows. First, in order to create a spot array, the unit element is replicated and overlapped with a certain pitch T . The total field distribution u_{D} of the Direct-imaging DLA can thus be calculated by the following equation

$$u_{\text{D}} = \sum_{m,n} u_{\text{unit}}(x - mT, y - nT). \quad (3.18)$$

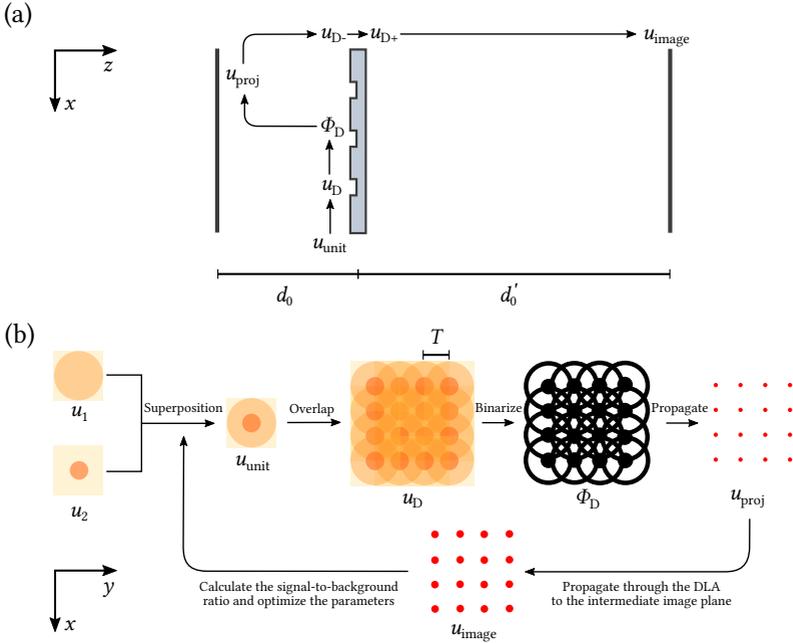


Figure 3.10: Design process of the Direct-imaging DLAs. (a) Pipeline for the image simulation. (b) Flow chart for the design and optimization.

It is worth noting that for the Direct-imaging DLAs, the overlapping has to be done with the unit cells on the DLA plane, instead of the spot array on the object plane. This is to provide a one-to-one correspondence between the projected spot and the intermediate image spot. Then, in order to simulate a binary phase mask, the phase of the DLA pattern is also extracted and binarized in the same way as Equation 3.4, which is given by

$$\phi_D = \text{mod} \left(\left\lfloor \frac{\arg(u_D)}{\pi} + B \right\rfloor, 2 \right) \pi. \quad (3.19)$$

Next, similar to the See-through DLAs, the simulation of the imaging process begins with the projection of the spot array with plane wave illumination. The projected field is represented as

$$u_{\text{proj}} = \mathcal{S}\{e^{-i\phi_D}, k_0, d_0\}. \quad (3.20)$$

Once again, the projected spots are reflected by a mirror surface, and the field propagates back to the DLA plane. The field right before the DLA plane is also expressed as

$$u_{D-} = \mathcal{S}\{u_{\text{proj}}, k_0, d_0\}. \quad (3.21)$$

Furthermore, the Direct-imaging DLA is also regarded as a thin element and its thickness is neglected. Then, the field right after the DLA plane is expressed as

$$u_{D+} = u_{D-} \cdot e^{-i\phi_D}. \quad (3.22)$$

Finally, the field on the intermediate image plane can be calculated by the propagation of u_{D+} through a designed distance of d'_0 , which is given by

$$u_{\text{image}} = \mathcal{S}\{u_{D+}, k_0, d'_0\}. \quad (3.23)$$

Afterwards, the intermediate image is collected by the objective onto the camera sensor. However, unlike the See-through DLAs, it is not necessary to further include such a process in this case, since there is no additional disturbance in between, and the intermediate image already reflects the quality and the SNR of the final image.

3.2.3 Simulation and optimization

The parameters for the Direct-imaging DLA are also optimized according to the PBR of the unit cell, which is defined in Equation 3.11. Since there is no imaging objective with aberrations involved in the simulation, all the $N \times N$ unit cells are counted in the calculation of the average PBR of the intermediate image.

$$\text{PBR}_N = \sum_{m,n}^N \frac{\text{PBR}(m,n)}{N^2}, \quad (3.24)$$

Similarly, the parameters need to be optimized iteratively in every loop in order to maximize the average PBR. Again, some of the parameters are defined according to the specifications, including the wave number, the number of the spots, and the pitch of the array T , while the others are treated as variables to be optimized, including the working distance d_0 , the distance to the intermediate image d'_0 , the ratio between the two lens components of the unit element W_D and the binarization factor B . The optimization argument can also be expressed by

$$(B, d_0, d'_0, W_D)_{\text{opt}} = \arg \max_{B, d_0, d'_0, W_D} \text{PBR}_N(B, d_0, d'_0, W_D). \quad (3.25)$$

Comparably, the four parameters are also divided into two groups, among which W_D is independent while d_0 , d'_0 and B are coupled. The reason for the coupling may come from the superposition of different field distributions. Consequently, constructive and destructive interference will occur during the superposition, which is influenced by these parameters. It is natural to come into a conclusion that more destructive interference can lead to a weaker signal. For example, if two field distributions cancel each other completely, the resulting phase mask will simply become a piece of glass, and no desired signal will be produced. However, such superposition is complicated to quantify, and thus the strategy for the optimization is to simply make a scan of the coupled parameters and to choose the best group of them.

The optimization process of the Direct-imaging DLAs is shown in Figure 3.11, which is generally the same as that of the See-through DLAs. First, as previously stated, the initial values are set for the parameters, for example, $d_0 = 1$ mm, $d'_0 = 20$ mm, $B = 0.5$ and $W_D = 20$. Then, the process follows the routine to optimize W_D and the other three variables separately. This time, the three variables are scanned within the defined region in order to achieve the maximum average PBR. Finally, a gradient descent optimization is used for the fine tuning of the parameters.

As an example, a Direct-imaging DLA pattern is optimized to generate a 5×5 spot array with a pitch of $100 \mu\text{m}$. The simulation wavelength is 785 nm,

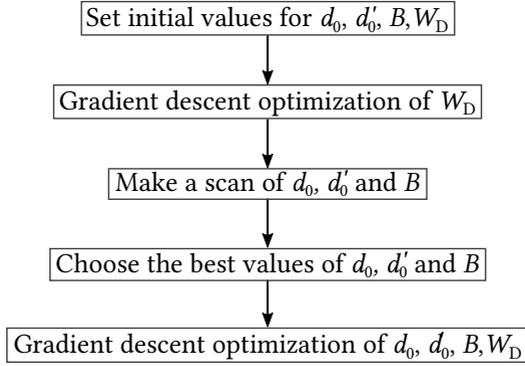


Figure 3.11: Optimization strategy for the Direct-imaging DLAs.

the feature size is $0.4\ \mu\text{m}$ and the size of the unit cell is $2\ \text{mm}$. The other parameters are optimized iteratively as described above with $d_0 = 1.11\ \text{mm}$, $d'_0 = 21.26\ \text{mm}$, $B = 0.98$ and $W_D = 15.38$. The simulated intermediate image is shown in Figure 3.12 (a), while the cross section of the central spot in the array at $y = 0$ is shown in Figure 3.12 (b). There appear to be some deviations of the brightness of individual spots, which is related to the overlapping and binarization of the fields. If required, such deviations can be compensated by applying different W_D for different unit cells. Nevertheless, it is clearly shown that all the 5×5 spots are visible in the intermediate image. However, with the increasing number of the spots, the intermediate image become noisy with irregular interference patterns. For a 9×9 array in Figure 3.12 (c), the intermediate image shows significantly more noise. For an 11×11 array in Figure 3.12 (e), the spots become almost indistinguishable.

Such noise is mainly caused by the cross-talks due to the superposition of different fields and overlapping apertures. There are several kinds of cross-talks. On the one hand, as previously mentioned, the unit cell of a Direct-imaging DLA is composed of two components which are two kinds of lenses with different focal lengths. They act as the illumination lens and the imaging lens respectively. However, they do not work separately for the illumination and the imaging as required. Instead, they always take effect at the same time.

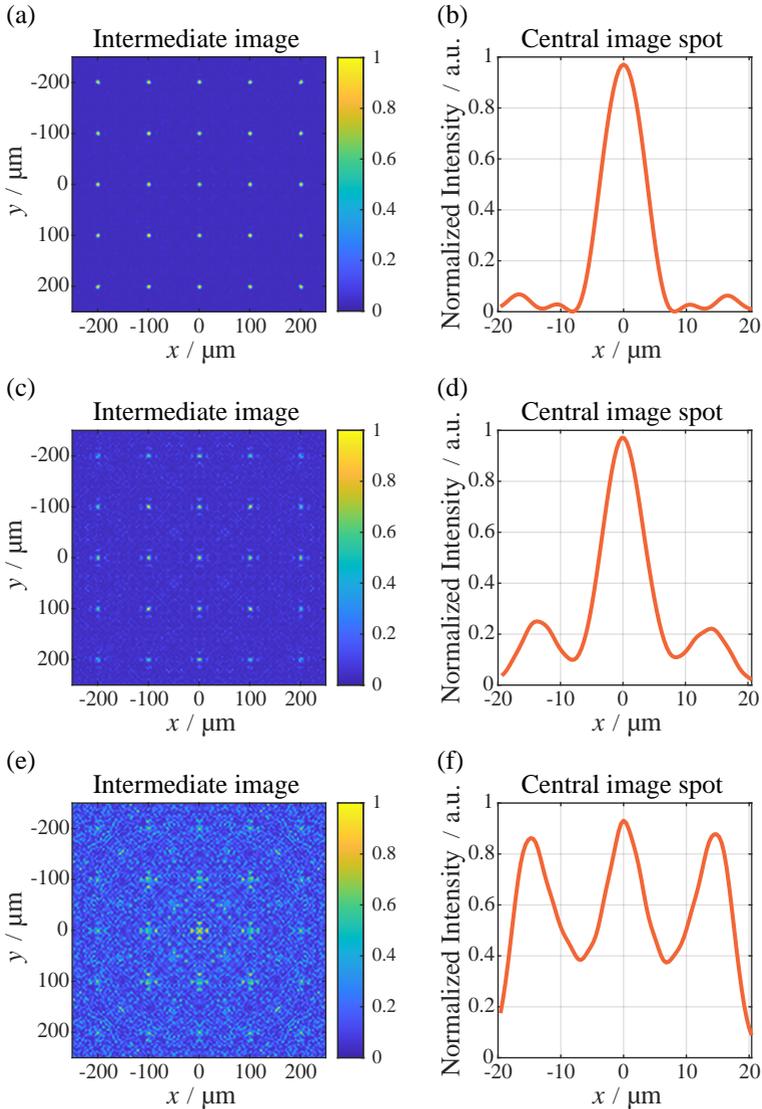


Figure 3.12: Simulation of spots in the intermediate image. (a) A 5×5 spot array. (b) The central spot of the 5×5 array. (c) A 9×9 spot array. (d) The central spot of the 9×9 array. (e) An 11×11 spot array. (f) The central spot of the 11×11 array.

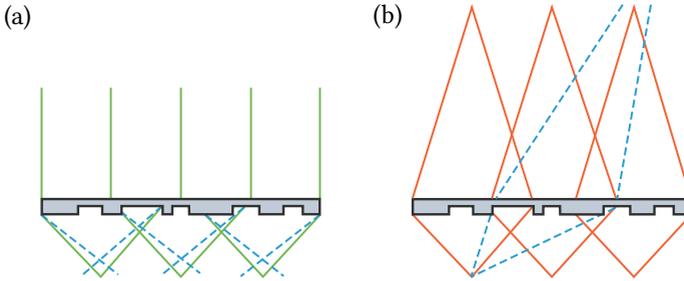


Figure 3.13: Disturbance which causes irregular interference patterns. (a) Out-of-focus spots produced the superimposed lenses. (b) Imaging through the adjacent lenses.

When the illumination spots are projected, the imaging lenses also produce blurred spots around the desired spots as shown in Figure 3.13 (a).

On the other hand, in the imaging path, due to the overlapping apertures, one spot not only passes through the designed lens to form a spot on the intermediate image, but it also passes through the adjacent lenses to form an extra blurred spot. As shown in Figure 3.13 (b), the blurred spot interferes with the original spot and adds noise to the image. Furthermore, a binary phase mask has a diffraction efficiency lower than 50 percent in theory [Swa89]. Other diffraction orders also enter the imaging system and become part of the background noise. Consequently, as the number of spots increases, the SNR of the image decreases.

The main cause for the above-mentioned cross-talks is too much overlapping of the unit elements. The overlapping can be reduced by decrease of the working distance, decrease of the NA, or increase of the pitch of the elements. One possible solution to reduce the overlapping can be the arrangement of the spots in a line instead of a 2D array. In this case, there is no interference from the other dimension, e.g., from the top and bottom spots. Meanwhile, the spots farther away will contribute less disturbance. Thus, the overall disturbance for a single spot can be controlled in an acceptable level and the line of spots can be extended infinitely. Fig. 3.14 shows a line of 27 spots with a pitch of $120\ \mu\text{m}$. The feature size, the wavelength and the working distances are the same as the simulations in Figure 3.13. As shown in the picture, the central spot which

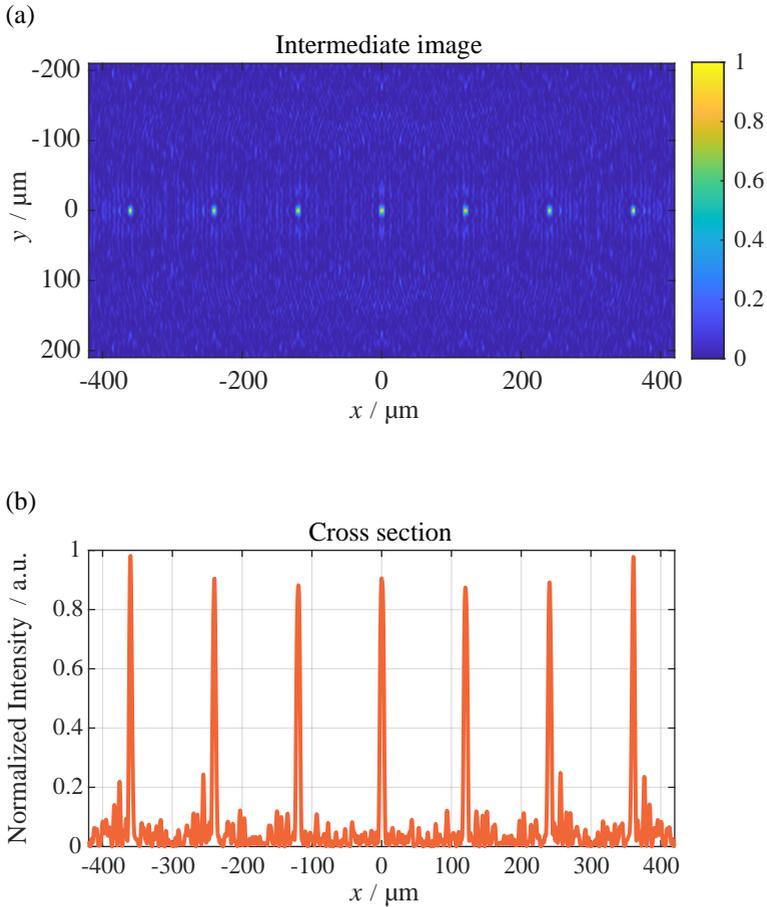


Figure 3.14: Simulation of a Direct-imaging DLA which produces a line of 27 spots with a pitch of 120 μm . (a) The central 9 spots on the intermediate image. (b) Cross section of the spots at $y = 0$.

receives the most disturbance from other spots is still clearly visible. By further extension of the spot line, the disturbance from the unit cells far away from the center should become negligible, which makes it possible for a line sensor.

In summary, in this chapter, the idea of superposition is proposed to overcome the limit of previous DLAs and to enable their application in opaque surface measurement. Two different DLA concepts are proposed and simulated. The first one is the See-through design. It increases the zero-order diffraction efficiency and reduces the disturbance of the fields to let the illumination spots be imaged through it. Thus, it can be adapted to the reflected-light microscope setup. It provides a good SNR and contrast of the image but it is limited in the axial sensitivity for surface measurement. The second one is the Direct-imaging design, which has the same depth discerning capability as the traditional high-NA objectives and provides excellent resolution in all directions. However, it suffers from a low SNR of the image due to the overlapping of different wave components. Such side effects can be relieved by the arrangement of the spots in a sparse grid or a line. According to the simulation and design process, prototypes of the DLAs are produced. Different experiments are carried out and the properties of the DLAs are investigated in detail. The results are shown and discussed in the next chapter.

4 Experiment Results

In this chapter, the experiment results of the DLAs are presented. Following the simulation methods in Chapter 2 and the design procedures in Chapter 3, two binary phase masks are produced by electron-beam lithography. The first one is designed for a wavelength of 785 nm with a minimum feature size of 0.4 μm and a glass thickness of 6.25 mm. The second one is designed for a wavelength of 488 nm with a minimum feature size of 0.25 μm and a glass thickness of 1 mm, which is mainly used for fluorescence microscopy. The two DLA prototypes are denoted as the DOE-785 and the DOE-488 in the following contents. There are a number of DLA patterns with different parameters on both prototypes. Different kinds of experiments are conducted to test the performance of the DLAs.

In Section 4.1, the spots produced by the See-through and the Direct-imaging DLAs are measured, and the results are compared with the simulations. In Section 4.2, a resolution target is scanned by the proposed DLA-based confocal microscopes with low-NA objectives to demonstrate the lateral resolution enhancement by the DLAs. Then, in Section 4.3, a step height calibration target is measured by the DLAs to examine their axial sensitivity for surface measurement. Particularly, interference is utilized to improve the axial measurement sensitivity of the See-through DLAs. At last, in Section 4.4, fluorescent samples are measured to show the application of the See-through DLAs for high-resolution large-area fluorescence microscopy.

4.1 Spot characterization

In this section, the spots produced by both kinds of the DLAs are measured. The size of the spots is directly related to the resolution of the DLA-based confocal microscope setups. Besides, it is one of the most straightforward ways to examine the accuracy of the simulations.

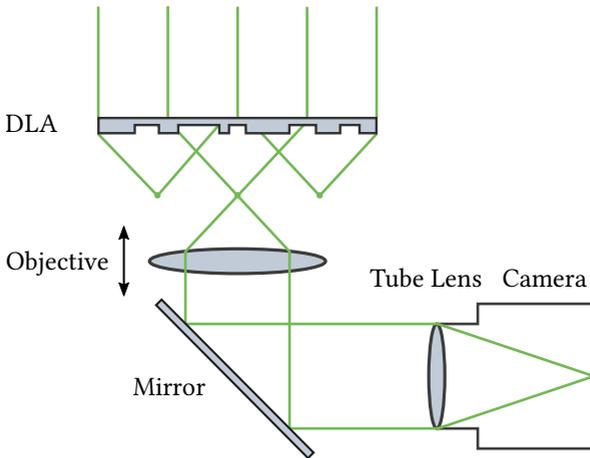


Figure 4.1: Schematic of the setup for the spot measurement.

The spot measurement setup is shown in Figure 4.1. A point laser source is collimated by an objective. The collimated light illuminates a DLA which produces a spot array. The spots are imaged by a 150×0.9 NA objective (Leica HCX PL APO $150\times/0.9$) and a 200 mm tube lens onto the camera sensor (Andor Zyla 5.5) with a pixel size of $6.5\ \mu\text{m}$. The objective under the DLA is moved axially by a vertical stage (Physik Instrumente L-306) to measure the 3D shape of the spots.

It is worth noting that the NA of the imaging objective in Figure 4.1 needs to be larger than that of the spot to be measured. In this case, the imaging process is not the convolution between the intensity of the spot and the intensity PSF of the objective. Instead, it can be treated as that the aperture of the imaging

objective is only partly filled by the light from the spot. Therefore, the size of the spot on the camera sensor reflects the actual size of the spot produced by the DLA. Both the See-through DLAs and the Direct-imaging DLAs are measured in this way. The experiment results are shown in the following sections.

4.1.1 Spot measurement of See-through DLAs

4.1.1.1 See-through DLAs at a wavelength of 785 nm

First, a See-through DLA pattern in the DOE-785 with a feature size of $0.4\ \mu\text{m}$ is measured by the above setup in Figure 4.1. A single-frequency fiber coupled diode laser (Thorlabs LP785-SAV50) is collimated to illuminate the DLA.

The selected See-through DLA produces an 11×11 spot array with a pitch of $100\ \mu\text{m}$ and a working distance of $1095\ \mu\text{m}$. The total area of the DLA on the phase mask is $3\ \text{mm} \times 3\ \text{mm}$. It has a plane-wave component weight $W_S = 2$ and a binarization factor $B = 0.36$. The objective in Figure 4.1 is scanned in the axial direction to measure the 3D shape of the spot.

The measurement of the spot in the center of the array is shown in Figure 4.2. The focal plane is defined by the pixel with the maximum intensity in the measurement. The measured lateral and axial shapes of the spot can be found in Figures 4.2 (a) and (c). The phase pattern has a pixel size of $0.4\ \mu\text{m}$ and the lateral simulation result is fitted to an Airy disc for comparison with the measurements, which is shown in Figure 4.2 (b) and (d). The simulated lateral FWHM is $0.496\ \mu\text{m}$ while the measured lateral FWHM is $0.516\ \mu\text{m}$, which corresponds to an NA of 0.78 according to Equation 2.93. The measured lateral spot size is close to, yet slightly smaller than the simulation result. The difference can be accounted for several reasons, such as the imperfect collimation, the errors in the sizes of the camera pixels, the deviation of the actual laser wavelength to the designed wavelength, the mechanical tolerances in the measurement setup and the inaccuracy caused by the scalar diffraction theory. In the axial direction, the simulated FWHM is $2.432\ \mu\text{m}$ while the measured lateral FWHM is $2.444\ \mu\text{m}$, which are also close to each other.

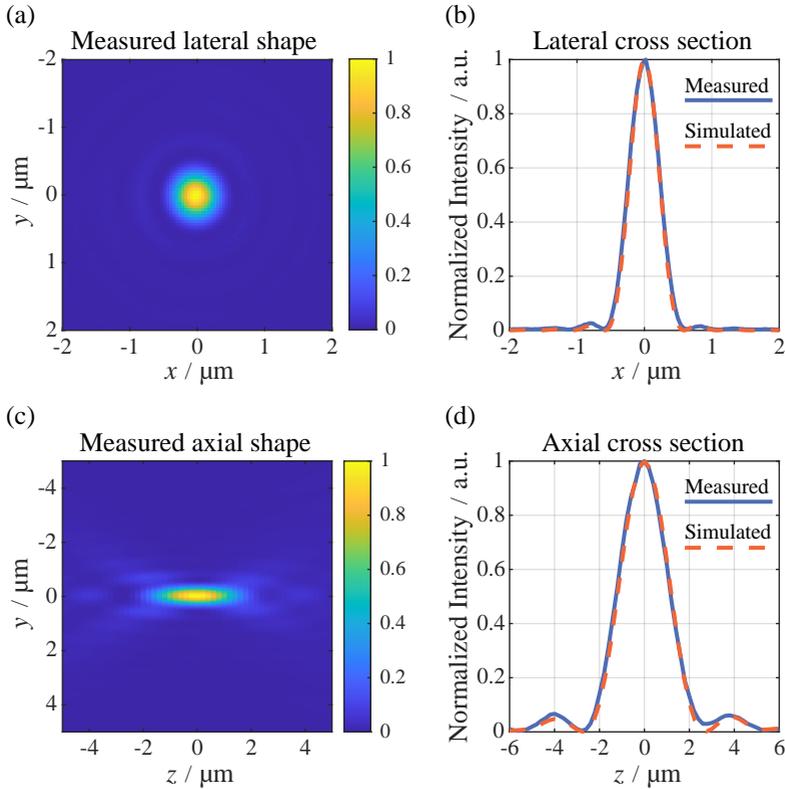


Figure 4.2: Measurement of the central spot in an 11×11 array with a pitch of $100 \mu\text{m}$ projected by a See-through DLA designed for a wavelength of 785 nm . The color bar represents the normalized intensity. (a) Lateral shape of the spot at $z=0$. (b) Lateral cross section of the spot at $z=0, y=0$ with a simulated FWHM of $0.496 \mu\text{m}$, and a measured FWHM of $0.516 \mu\text{m}$. (c) Axial shape of the produced spot at $x=0$. (d) Axial cross section of the spot at $x=0, y=0$ with a simulated FWHM of $2.432 \mu\text{m}$, and a measured FWHM of $2.444 \mu\text{m}$.

Furthermore, the spot in the corner of the 11×11 array for the same See-through DLA pattern is shown in Figure 4.3. The simulation and experiment results also fit well with each other. In the lateral direction, the simulated FWHM of the spot is $0.520 \mu\text{m}$ and the measured FWHM is $0.529 \mu\text{m}$, which corresponds to an NA of 0.76. In the axial direction, the simulated FWHM

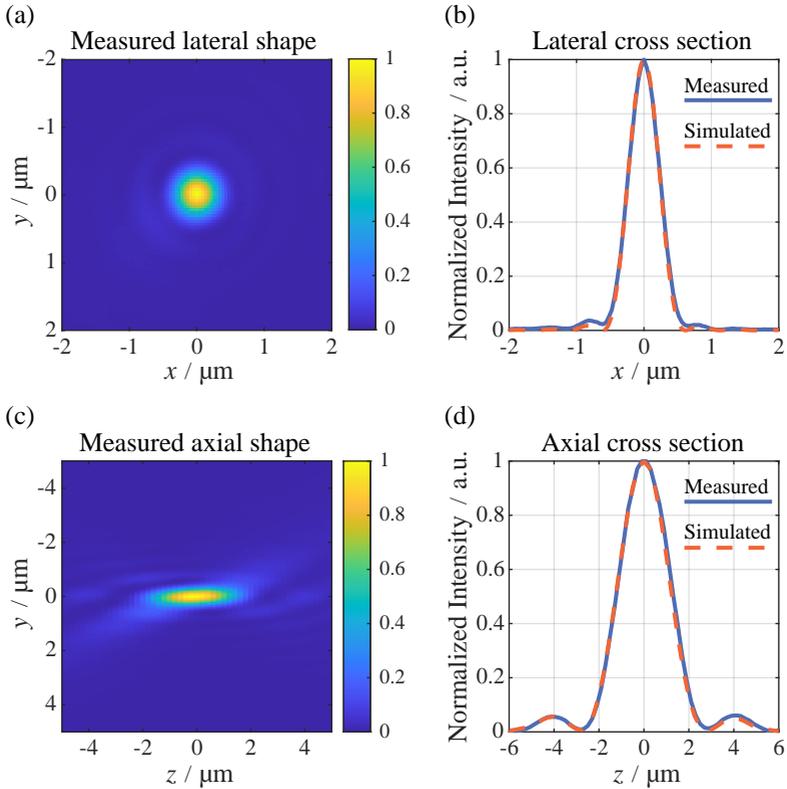


Figure 4.3: Measurement of the spot in the corner of an 11×11 array with a pitch of $100 \mu\text{m}$ projected by a See-through DLA designed for a wavelength of 785 nm . The color bar represents the normalized intensity. (a) Lateral shape of the spot at $z=0$. (b) Lateral cross section of the spot at $z=0, y=0$ with a simulated FWHM of $0.520 \mu\text{m}$, and a measured FWHM of $0.529 \mu\text{m}$. (c) Axial shape of the produced spot at $x=0$. (d) Axial cross section of the spot at $x=0, y=0$ with a simulated FWHM of $2.527 \mu\text{m}$, and a measured FWHM of $2.552 \mu\text{m}$.

is $2.527 \mu\text{m}$ and the measured FWHM is $2.552 \mu\text{m}$. It is noted that the spot in the corner becomes slightly larger and the axial shape of the spot exhibits certain asymmetry. This is mainly due to the fact that the spot does not receive contribution from other unit cells symmetrically since it is located in the corner of the pattern. Besides, on the edge of a collimated beam, the field

may also differ from a plane wave, which makes the spot look asymmetrical in the axial direction.

4.1.1.2 Original DLAs at a wavelength of 785 nm

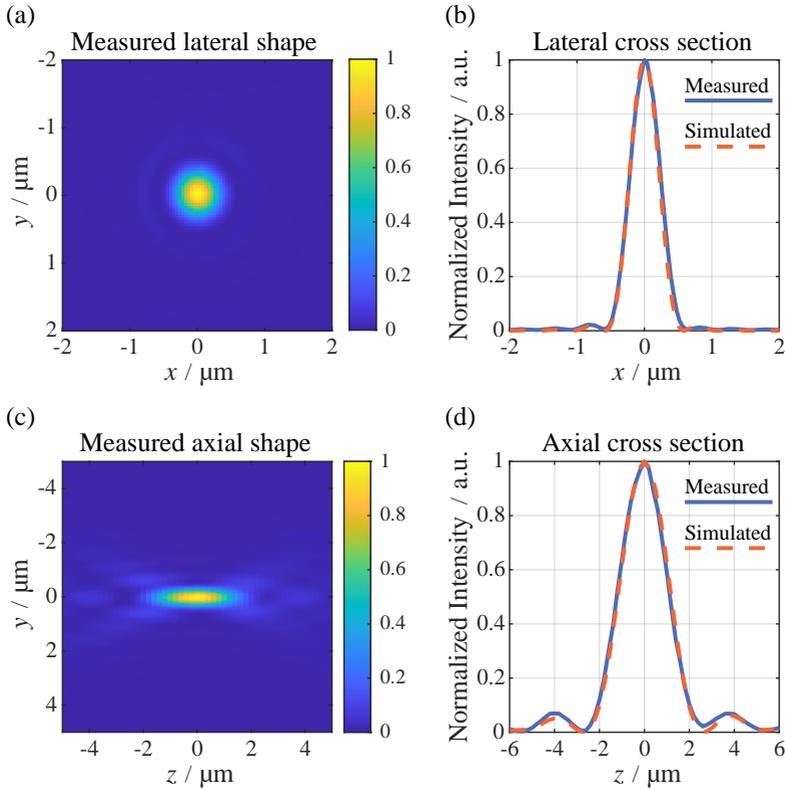


Figure 4.4: Measurement of the spot in the center of an 11×11 array with a pitch of $100 \mu\text{m}$ projected by an original DLA designed for a wavelength of 785 nm . The color bar represents the normalized intensity. (a) Lateral shape of the spot at $z=0, y=0$. (b) Lateral cross section of the spot at $z=0, y=0$ with a simulated FWHM of $0.496 \mu\text{m}$, and a measured FWHM of $0.511 \mu\text{m}$. (c) Axial shape of the produced spot at $x=0$. (d) Axial cross section of the spot at $x=0, y=0$ with a simulated FWHM of $2.384 \mu\text{m}$, and a measured FWHM of $2.372 \mu\text{m}$.

An original DLA without the plane-wave component is also measured to check whether the plane-wave component affects the spot quality. The original DLA pattern has the same parameters as the See-through DLA above except that it has a different binarization factor $B = 0.41$ and a zero W_S . The measurement of the central spot in the 11×11 array are shown in Figure 4.4. The measured spot has a lateral FWHM of $0.511 \mu\text{m}$ and an axial FWHM of $2.372 \mu\text{m}$. Compared to the results in Figure 4.2, the spot size of the See-through DLA is a bit larger than that of the original DLA, which indicates that the plane-wave component affects the spot quality within a certain range. An increase of the plane-wave component will further increase the spot size. For example, for another See-through pattern with $W_S = 13$, its central spot is measured with a lateral FWHM of $0.601 \mu\text{m}$ and an axial FWHM of $2.533 \mu\text{m}$, which corresponds to an NA of 0.67. However, such a side effect can be controlled to a negligible level by comparing the results in Figure 4.2 and 4.4.

4.1.1.3 See-through DLAs at a wavelength of 488 nm

Moreover, a See-through DLA designed with higher resolution in the DOE-488 is also measured. The experiment setup is the same as above except that the wavelength of the laser (Integrated Optics MatchBox) is changed to 488 nm. It produces an 13×13 spot array with a pitch of $80 \mu\text{m}$, a working distance of $501 \mu\text{m}$, a total area of $3 \text{ mm} \times 3 \text{ mm}$ and a minimum feature size of $0.25 \mu\text{m}$. The measurement of the central spot in the array is shown in Figure 4.5. The measured lateral FWHM of the spot is $0.304 \mu\text{m}$, which corresponds to an NA of 0.83. By simulation, the lateral FWHM is $0.292 \mu\text{m}$, which is close to the measurement. In the axial direction, the measured FWHM is $0.954 \mu\text{m}$ while the simulated result is $0.823 \mu\text{m}$.

The spot in the corner of the array is also measured, which is shown in Figure 4.6. In the lateral direction, the measured FWHM of the spot is $0.340 \mu\text{m}$, which corresponds to an NA of 0.74, and the simulated FWHM is $0.325 \mu\text{m}$. In the axial direction, the measured FWHM is $0.881 \mu\text{m}$ and the simulated FWHM is $0.757 \mu\text{m}$. The difference between the simulation and the measurement is larger than that of the DLA designed for the 785 nm wavelength. The reason

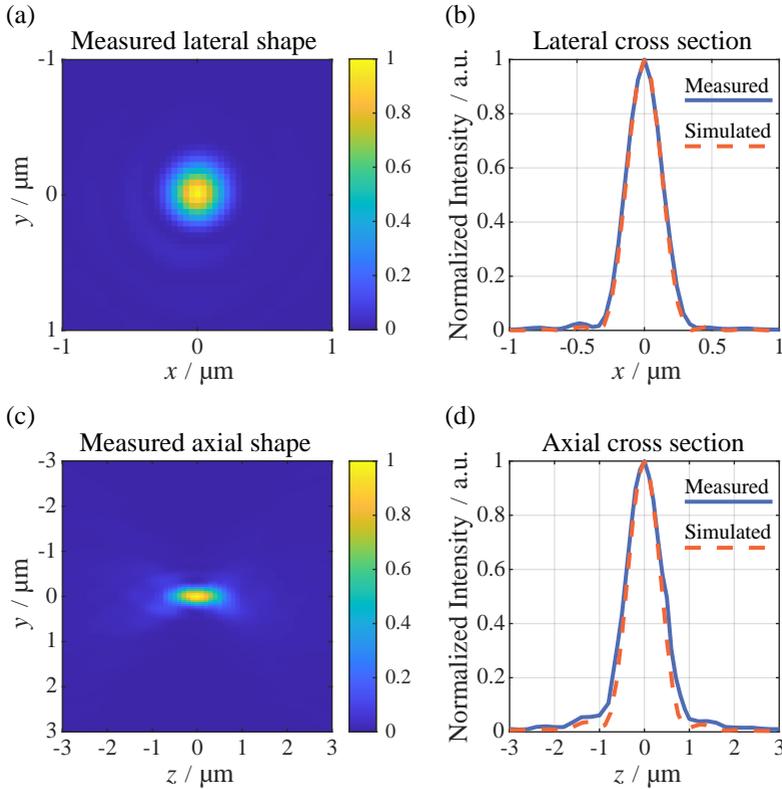


Figure 4.5: Measurement of the central spot in a 13×13 array with a pitch of $80 \mu\text{m}$ projected by a See-through DLA designed for a wavelength of 488 nm . The color bar represents the normalized intensity. (a) Lateral shape of the spot at $z=0$. (b) Lateral cross section of the spot at $z=0, y=0$ with a simulated FWHM of $0.292 \mu\text{m}$, and a measured FWHM of $0.304 \mu\text{m}$. (c) Axial shape of the produced spot at $x=0$. (d) Axial cross section of the spot at $x=0, y=0$ with a simulated FWHM of $0.823 \mu\text{m}$, and a measured FWHM of $0.954 \mu\text{m}$.

can be that the DLA for a shorter wavelength and a higher NA is more sensitive to the mechanical tolerances in the measurement setup, the aberrations of the collimated beam, and the fabrication errors of the phase mask during the lithography process. Besides, the laser has a bandwidth of around 1 nm , which also affects the spot size due to chromatic aberrations.

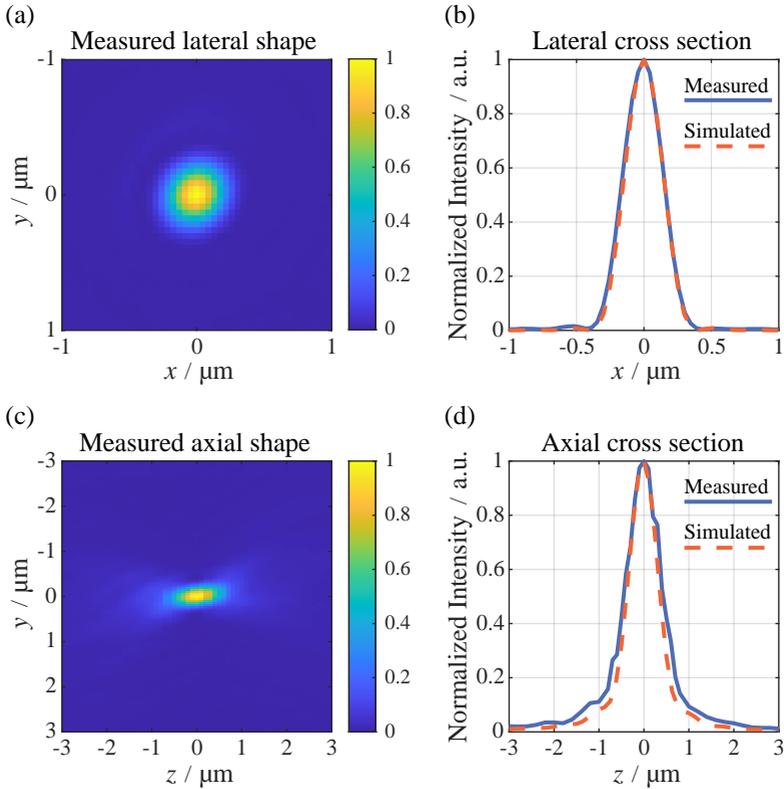


Figure 4.6: Measurement of the spot in the corner of a 13×13 array with a pitch of $80 \mu\text{m}$ projected by a See-through DLA designed for a wavelength of 488 nm . The color bar represents the normalized intensity. (a) Lateral shape of the spot at $z=0$. (b) Lateral cross section of the spot at $z=0, y=0$ with a simulated FWHM of $0.325 \mu\text{m}$, and a measured FWHM of $0.340 \mu\text{m}$. (c) Axial shape of the produced spot at $x=0$. (d) Axial cross section of the spot at $x=0, y=0$ with a simulated FWHM of $0.757 \mu\text{m}$, and a measured FWHM of $0.881 \mu\text{m}$.

It is worthwhile to mention that during the design process of the See-through DLA pattern in Figure 4.5, the size of the initial spot is about only a single pixel, which is $0.25 \mu\text{m}$. Besides, geometrically, the working distance of $501 \mu\text{m}$ and the DLA aperture size of 3 mm allow a maximum NA of 0.95 . However, such a design target is not achieved in both the simulation and the measurement.

This is mainly due to the limitation of the minimum feature size of the phase mask. To further increase the NA of the spots, the phase mask needs to have an even smaller feature size, which significantly increases the manufacturing cost. Besides, for such a high NA, the scalar diffraction theory also differs more from the rigorous solutions of the Maxwell's equations, which places the limitation on the spot size produced by a DLA. Nevertheless, to the best of the author's knowledge, the experiment in this work has so far shown the highest NA and the smallest spot size by the array illuminators, which lays the foundation for high-resolution large-area measurement with the proposed microscope setups.

4.1.2 Spot measurement of Direct-imaging DLAs

The spots produced by the Direct-imaging DLAs are also tested by the same setup. The selected DLA pattern is designed for a wavelength of 785 nm, which produces a 5×5 spot array with a pitch of 100 μm, a working distance of 1110 μm and a total area of 2.4 mm × 2.4 mm on the phase mask. It has a lens ratio $W_D = 15.4$ and a binarization factor $B = 0.98$.

The measurement of the central spot is shown in Figure 4.7. The measured lateral FWHM is 0.573 μm and the axial FWHM is 2.612 μm, which corresponds to an NA of 0.7. The measurements are also close to the simulation results with slight deviations. The axial cross section shows asymmetry and has a

Table 4.1: Measurement of the spots produced by the DLAs.

	Position	Measurement / μm		Simulation / μm	
		LFWHM	AFWHM	LFWHM	AFWHM
See-through DLA at 785 nm	Center	0.516	2.444	0.496	2.432
	Corner	0.529	2.552	0.520	2.527
Original DLA at 785 nm	Center	0.511	2.372	0.496	2.384
	Corner	0.532	2.407	0.530	2.382
See-through DLA at 488 nm	Center	0.304	0.954	0.292	0.823
	Corner	0.340	0.881	0.325	0.757
Direct-imaging DLA at 785 nm	Center	0.573	2.612	0.562	2.664
	Corner	0.560	2.584	0.554	2.592

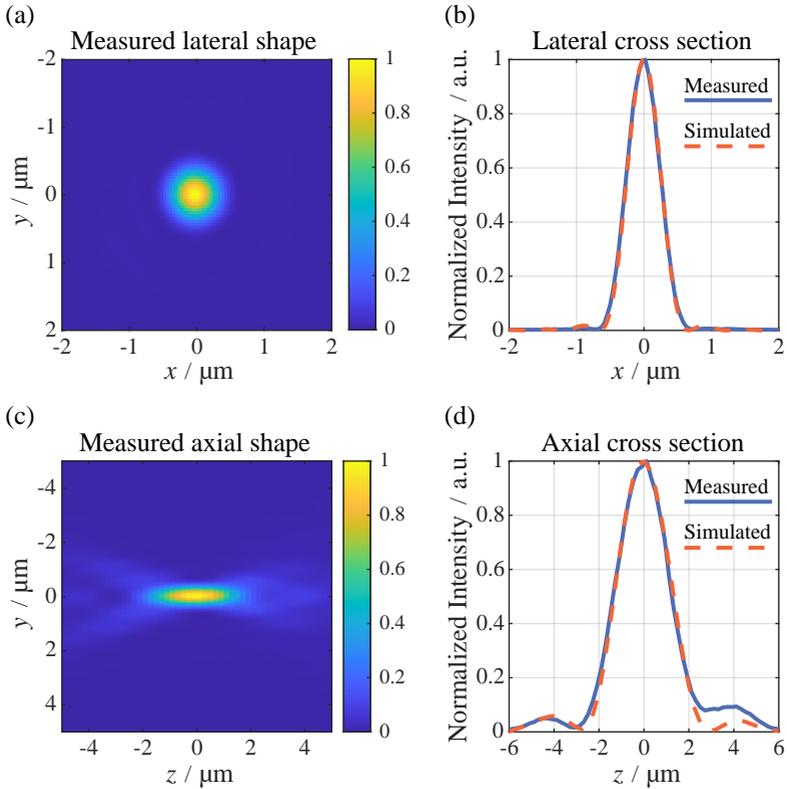


Figure 4.7: Measurement of the central spot in a 5×5 array with a pitch of $100 \mu\text{m}$ projected by a Direct-imaging DLA designed for a wavelength of 785 nm . The color bar represents the normalized intensity. (a) Lateral shape of the spot at $z=0$. (b) Lateral cross section of the spot at $z=0, y=0$ with a simulated FWHM of $0.562 \mu\text{m}$, and a measured FWHM of $0.573 \mu\text{m}$. (c) Axial shape of the produced spot at $x=0$. (d) Axial cross section of the spot at $x=0, y=0$ with a simulated FWHM of $2.664 \mu\text{m}$, and a measured FWHM of $2.612 \mu\text{m}$.

smaller FWHM than the simulation possibly due to the imperfections in the collimated beam.

The measured spot in the corner of the array is shown in Figure 4.8. Unlike the See-through DLAs, the size of the spot in the corner is smaller than the one in the center in both experiments and simulations. Since the superposition

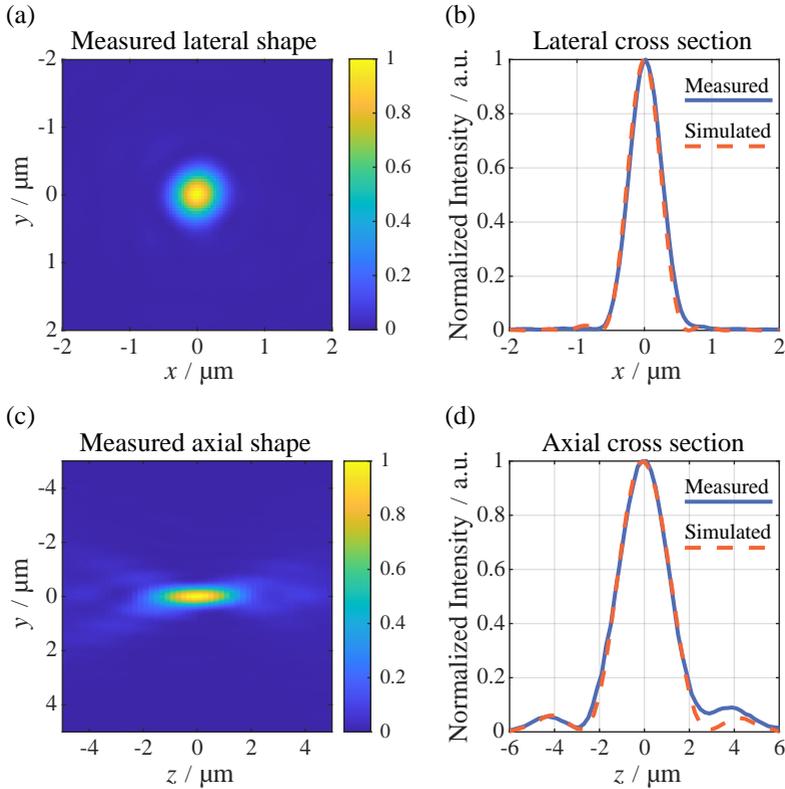


Figure 4.8: Measurement of the spot in the corner of a 5×5 array with a pitch of $100 \mu\text{m}$ projected by a Direct-imaging DLA designed for a wavelength of 785 nm . The color bar represents the normalized intensity. (a) Lateral shape of the spot at $z=0$. (b) Lateral cross section of the spot at $z=0, y=0$ with a simulated FWHM of $0.554 \mu\text{m}$, and a measured FWHM of $0.560 \mu\text{m}$. (c) Axial shape of the produced spot at $x=0$. (d) Axial cross section of the spot at $x=0, y=0$ with a simulated FWHM of $2.592 \mu\text{m}$, and a measured FWHM of $2.584 \mu\text{m}$.

in the Direct-imaging DLAs is not uniform, the central unit cell receives the most disturbance from the surrounding ones. Nevertheless, the measurements of the spot sizes still fit well with the simulations.

In conclusion, Table 4.1 summarizes the measurement and simulation results of the spots produced by the DLAs. In general, the measurements fit very well

with the simulations. The average difference in the lateral FWHM between the measurement and the simulation is 11 nm, and the average difference in the axial FWHM is 31 nm. It proves that the design and simulation framework in Chapter 3 produces reliable predictions for the performances of the DLAs.

4.2 Lateral measurement

After the characterization of the spots, the proposed DLAs are integrated in the confocal microscopes, and they are tested in various applications. In this section, Both the See-through and the Direct-imaging DLAs are used for lateral scanning of a resolution target. Combined with low-NA objectives, the resolution of the system is quantitatively measured by the spatial cut-off frequency. The setups and the experiment results are presented as follows.

4.2.1 Lateral measurement by See-through DLAs

The setup of the confocal microscope based on the See-through DLA is shown in Figure 4.9. A fiber-coupled laser passes through the relay lenses (Thorlabs AC254-075-A-ML) and it is again focused into a spot. The spot is placed on the back focal point of an objective. Light is reflected by a beam splitter and collimated by the objective. The collimated laser illuminates the DLA which produces a spot array. The spots illuminate the sample and they are again imaged by the objective and a 180 mm tube lens onto the camera (FLIR BFS-U3-28S5M-C), which has a pixel size of $4.5\ \mu\text{m}$. The central pixel of each spot on the image acts as a pinhole and its intensity is recorded at every scanning step. Subsequently, a confocal image is reconstructed by raster scanning of the sample by a piezo stage (Physik Instrumente P-616).

A USAF high resolution target (Newport HIGHRES-2) is used as the sample, whose highest spatial frequency is $3649\ \text{lp}\cdot\text{mm}^{-1}$. It is measured by the DLA-based confocal setup to test the lateral resolution of the system. The See-through DLA which has been measured in Figure 4.2 is evaluated in the setup. It produces an 11×11 spot array with an NA up to 0.78 and the spots are

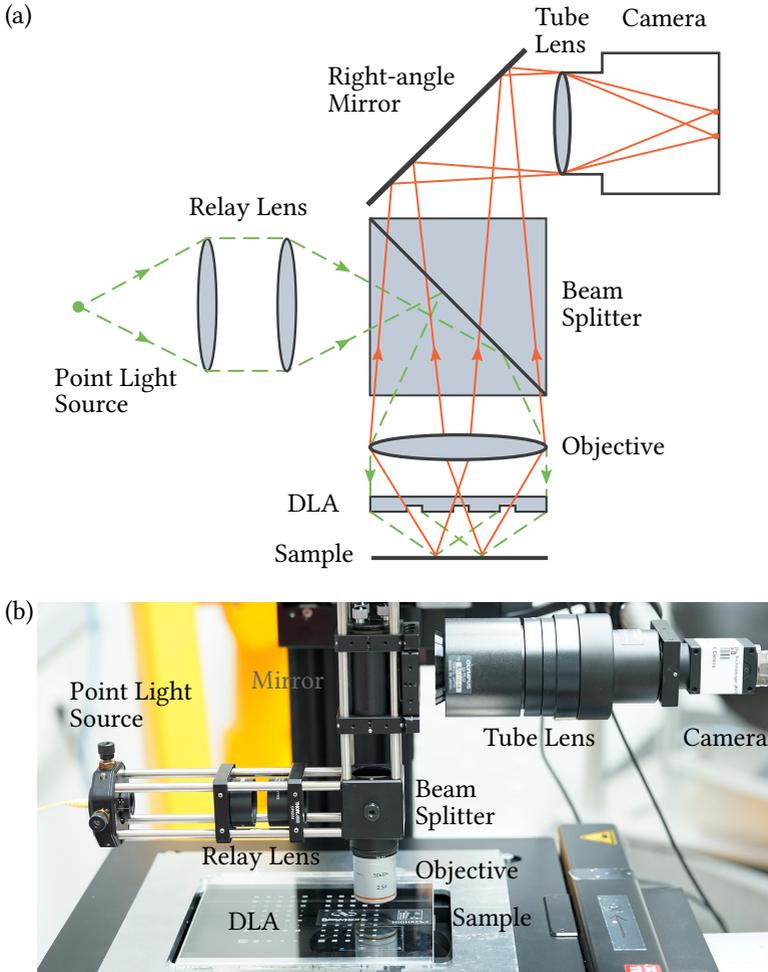


Figure 4.9: Measurement setup of a multi-spot confocal microscope based on the See-through DLA. The sample is scanned laterally and an image is reconstructed. (a) A schematic of the experiment setup. (b) A picture of the experiment setup.

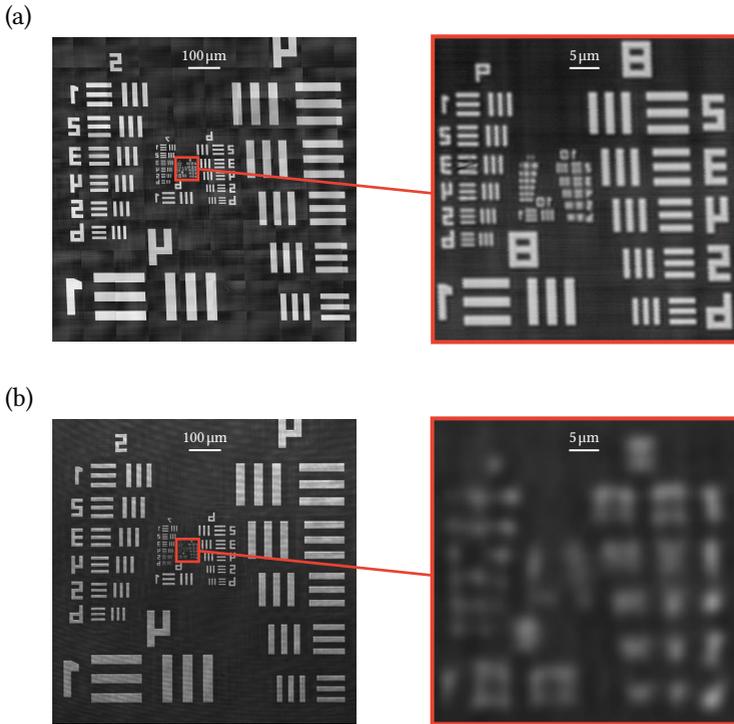


Figure 4.10: Images taken by a 5×0.15 NA objective at a wavelength of 785 nm. (a) Confocal scanning image by the See-through DLA with a scanning step of $0.2\ \mu\text{m}$. The maximum spatial frequency with non-zero contrast is $1024\ \text{lp}\cdot\text{mm}^{-1}$ of the element 10-1. (b) Wide-field image by the objective solely. The maximum spatial frequency with non-zero contrast is $203\ \text{lp}\cdot\text{mm}^{-1}$ of the element 7-5.

used for scanning the sample. A 785 nm laser (Thorlabs LP785-SF20) with a bandwidth of around 2 nm is used to decrease the noise caused by speckles.

First, a 5×0.15 NA objective (Nikon LU Plan Fluor $5\times/0.15$) is used in the setup to measure the sample. The resolution target is scanned by the piezo stage with a scanning step of $0.2\ \mu\text{m}$ and a camera frame rate of 100 fps. The result is shown in Figure 4.10 (a). The image shows that the measured highest spatial frequency with non-zero contrast is $1024\ \text{lp}\cdot\text{mm}^{-1}$ of the element 10-1, which corresponds to a line width of $0.488\ \mu\text{m}$. There are some stitching

artifacts and inclination in the image due to the non-orthogonality of the two axes of the piezo stage.

Theoretically, the resolution of the confocal setup based on the See-through DLA can be described by the intensity PSF in Equation 2.128. The modulation contrast with respect to the spatial frequencies is given by the optical transfer function (OTF) of the setup, which is the Fourier transform of the PSF. The amplitude of the OTF is also known as the modulation transfer function (MTF). It describes the contrast in the image at the corresponding spatial frequency. With a 0.78 NA DLA and a 0.15 NA objective, the theoretical MTF is calculated as the Figure 4.11 (a) shows. It can be seen that the modulation is close to zero around the spatial frequency of $1000 \text{ lp}\cdot\text{mm}^{-1}$, which agrees well with the spatial cut-off frequency of $1024 \text{ lp}\cdot\text{mm}^{-1}$ in the experiment. The theoretical modulation at $1024 \text{ lp}\cdot\text{mm}^{-1}$ is 0.24%. Besides, by comparison with the wide-field image taken solely by the same objective in Figure 4.10 with a measured spatial cut-off frequency of $203 \text{ lp}\cdot\text{mm}^{-1}$, the lateral resolution of the objective is enhanced by around 5 times by using the See-through DLA.

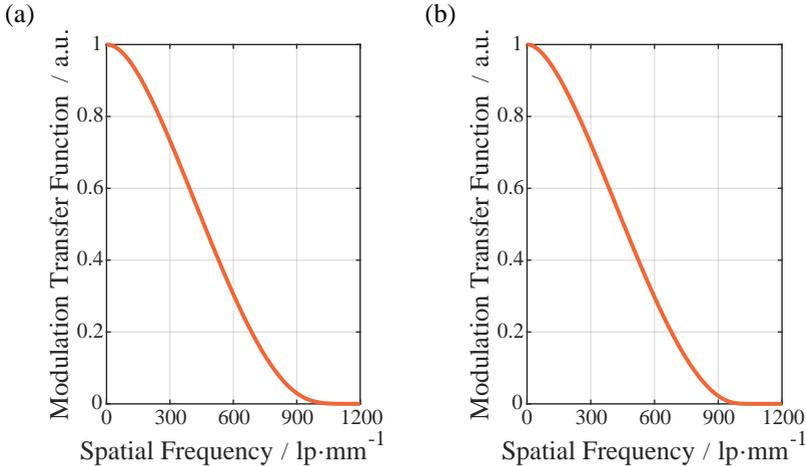


Figure 4.11: Simulated MTF of a confocal microscope based on a See-through DLA with 0.78 NA at a wavelength of 785 nm. (a) With a 0.15 NA objective. (b) With a 0.07 NA objective.

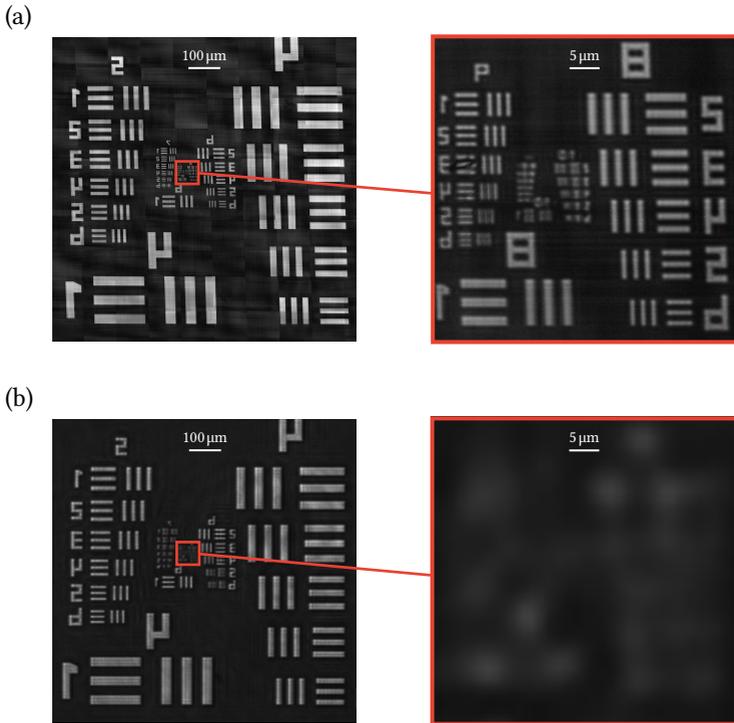


Figure 4.12: Images taken by a 2.5×0.07 NA objective at a wavelength of 785 nm. (a) Confocal scanning image by the See-through DLA with a scanning step of $0.2\ \mu\text{m}$. The maximum spatial frequency with non-zero contrast is $912\ \text{lp}\cdot\text{mm}^{-1}$ of the element 9-6. (b) Wide-field image by the objective solely. The maximum spatial frequency with non-zero contrast is $101\ \text{lp}\cdot\text{mm}^{-1}$ of the element 6-5.

Furthermore, a 2.5×0.07 NA objective (Leica FL PLAN $2.5\times/0.07$) is used to measure the resolution of the setup by an objective with an even lower NA. The target is also scanned with a step of $0.2\ \mu\text{m}$ and the measurement result is shown in Figure 4.12. The measured spatial cut-off frequency with a non-zero contrast is $912\ \text{lp}\cdot\text{mm}^{-1}$ of the element 9-6.

The theoretical MTF is shown in Figure 4.11 (b). It can be seen that although the NA of the objective is smaller than half of the previous one in Figure 4.11 (a), the overall MTF of the system only slightly decreases, which proves that the

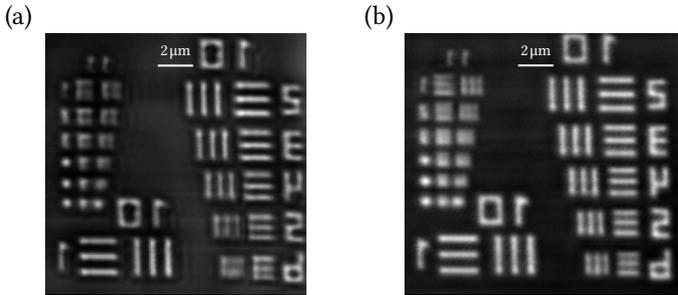


Figure 4.13: Confocal scanning images taken with a See-through DLA at a wavelength of 488 nm. (a) By a 5×0.15 NA objective, the maximum spatial frequency with non-zero contrast is $1625 \text{ lp}\cdot\text{mm}^{-1}$ of the element 10-5. (b) By a 10×0.3 NA objective, the maximum spatial frequency with non-zero contrast is $2048 \text{ lp}\cdot\text{mm}^{-1}$ of the element 11-1.

lateral resolution is mainly determined by the DLA and not by the objective. The calculated modulation is 0.02% at $1024 \text{ lp}\cdot\text{mm}^{-1}$, and thus the contrast of the stripes is too low to be visible. At $912 \text{ lp}\cdot\text{mm}^{-1}$, the calculated modulation is 1.73% , which is visible in the experiment. Moreover, the reduction in the spatial resolution is not solely attributed to the confocal imaging theory. Objectives with smaller NAs collect less light while the background noise almost remains the same. Therefore, the image taken by the $2.5\times$ objective has a lower SNR, which also leads to a reduction of the contrast and the spatial cut-off frequency. Nevertheless, compared with the wide-field image in Figure 4.12 (b) with a measured cut-off frequency of $101 \text{ lp}\cdot\text{mm}^{-1}$, the See-through DLA increases the lateral resolution by more than 9 times.

The See-through DLA pattern in DOE-488 with a higher resolution is also tested at a wavelength of 488 nm with the 5×0.15 NA objective and a 10×0.3 NA (Nikon LU Plan Fluor $10\times/0.3$) objective respectively. The spots produced by the DLA are measured in Figure 4.6 which have NAs up to 0.83. The central part of the resolution target in the measurement is shown in Figure 4.13 (a) and (b). The measured spatial cut-off frequencies are $1625 \text{ lp}\cdot\text{mm}^{-1}$ of the element 10-5 by the 0.15 NA objective, and $2048 \text{ lp}\cdot\text{mm}^{-1}$ of the element 11-1 by the 0.3 NA objective.

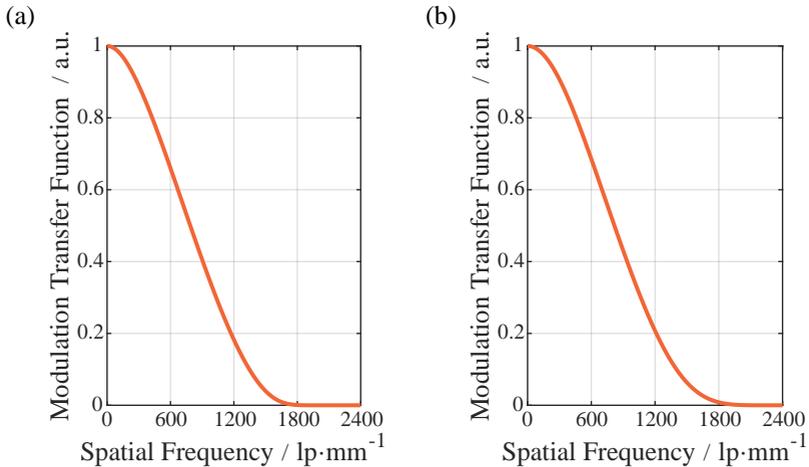


Figure 4.14: Simulated MTF of a confocal microscope based on a See-through DLA with 0.83 NA at a wavelength of 488 nm. (a) With a 0.15 NA objective. (b) With a 0.3 NA objective.

The theoretical MTF is shown in Figure 4.14. By a 0.15 NA objective, the modulation is 1% at $1625 \text{ lp}\cdot\text{mm}^{-1}$ of the element 10^{-5} , and 0.04% at $1825 \text{ lp}\cdot\text{mm}^{-1}$ of the next element 10^{-6} . By a 0.3 NA objective, the modulation at $2025 \text{ lp}\cdot\text{mm}^{-1}$ is 0.06%. The stripes are still visible due to the SNR of the image taken by a 0.3 NA objective is higher than that taken by a 0.15 NA objective. The SNR of the image is limited by the background noise, which mainly comes from the stray light. It overlaps with the confocal signals and composes a major part of the noise in the image, which will be discussed in detail in Section 4.3.4.

4.2.2 Lateral measurement by Direct-imaging DLAs

The schematic of the confocal scanning microscope setup based on the Direct-imaging DLA is shown in Figure 4.15. The setup is almost identical to that of the See-through DLA in Figure 4.9. The only difference is that the objective is lifted up, and the focal plane is placed on the intermediate image generated by the Direct-imaging DLA.

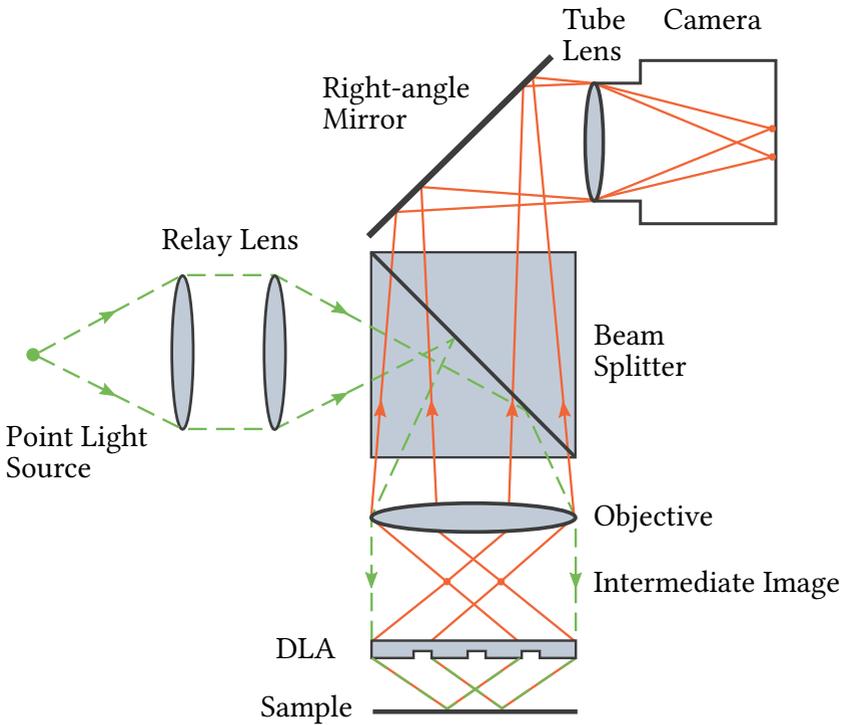


Figure 4.15: Schematic of a multi-spot confocal microscope setup based on the Direct-imaging DLA.

The same resolution target is scanned by the above setup with a scanning step of $0.2 \mu\text{m}$ and by the Direct-imaging DLA pattern tested in Section 4.1.2 which produces a 5×5 spot array with a pitch of $100 \mu\text{m}$. The imaging system takes images of the 5×5 intermediate spot array by the 5×0.15 NA objective. Similarly, the central pixel of each spot is used as the pinhole and the intensity is recorded at every scanning step to reconstruct the confocal image. The measured cut-off frequency is $1448 \text{lp}\cdot\text{mm}^{-1}$ of the element 10-4.

The Direct-imaging DLA should perform exactly as a high-NA finite-conjugate objective, which has the same NA of 0.7 as is measured in Figure 4.7. Theoretically, by Equation 2.128, the MTF of a confocal microscope with a 0.7 NA

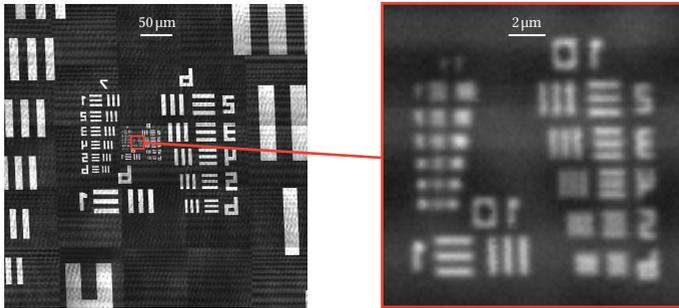


Figure 4.16: Confocal scanning image by the Direct-imaging DLA and the 5×0.15 NA objective with a scanning step of $0.2\ \mu\text{m}$ at a wavelength of $785\ \text{nm}$. The maximum spatial frequency with non-zero contrast is $1448\ \text{lp}\cdot\text{mm}^{-1}$ of the element 10-4.

objective is calculated in Figure 4.17, and the modulation at $1448\ \text{lp}\cdot\text{mm}^{-1}$ is 0.18%. Some stitching artifacts and inclination are observed as well in the image due to the non-orthogonality of the two axes of the piezo stage. There are also stripes due to the interference with the stray light, which becomes more apparent for the Direct-imaging DLAs, since the signal in the intermediate image is weaker compared to the See-through DLAs.

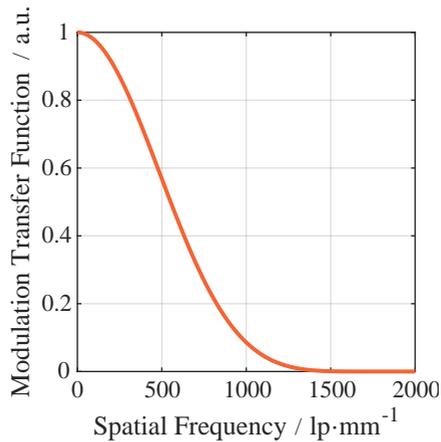


Figure 4.17: Simulated MTF of a confocal microscope based on an objective with $0.7\ \text{NA}$ at a wavelength of $785\ \text{nm}$.

In summary, the measured maximum spatial frequencies by different DLAs and objectives are listed in Table 4.2. The DLAs have remarkably enhanced the resolution of the objectives, which has been predicted by the theory in Section 2.2.3.

Table 4.2: Lateral measurement of the resolution target by the DLAs.

	Objective	Spatial cut-off frequency / lp·mm ⁻¹
See-through DLA at 785 nm	2.5× 0.07 NA	912
	5× 0.15 NA	1024
See-through DLA at 488 nm	5× 0.15 NA	1625
	10× 0.3 NA	2048
Direct-imaging DLA at 785 nm	5× 0.15 NA	1448

4.3 Axial measurement

In this section, the DLAs are used for surface measurement. The axial responses of the See-through DLAs and the Direct-imaging DLAs are characterized respectively. Besides, a new concept of using interference to increase the axial measurement sensitivity of the See-through DLAs is proposed. A calibrated step height target is used as the reference to test the measurement uncertainties of these methods.

4.3.1 Axial response of See-through DLAs

To characterize the axial response of the See-through DLAs, the same setup in Figure 4.9 is used. A plane mirror is used as the sample. The illumination spots are reflected by the mirror and they are imaged by the objective and the camera through the DLA. Then the mirror is scanned axially in z direction to measure the confocal axial response. The central pixel of the central spot in the 11×11 spot array on the image sensor is used as the pinhole. Its intensity is recorded when the mirror is moved. The 5× 0.15 NA objective is used for imaging.

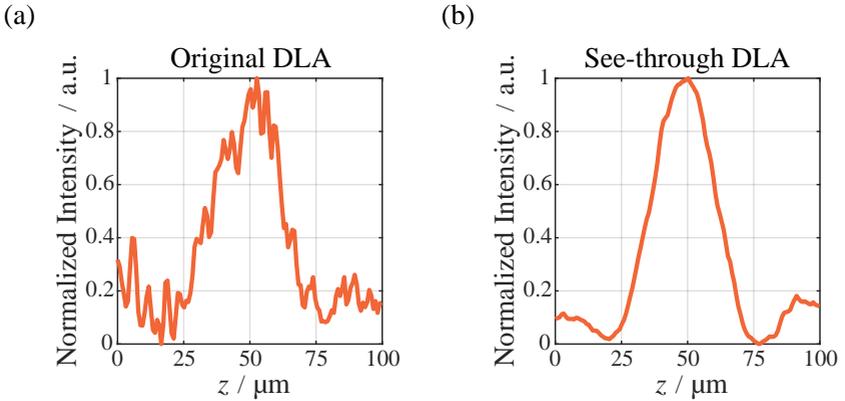


Figure 4.18: Confocal axial response by the 5×0.15 NA objective when a mirror is measured. (a) Axial response with the original DLA. (b) Axial response with the See-through DLA.

Figures 4.18 (a) and (b) show the results of the confocal axial responses for the original DLA and the See-through DLA respectively. The two signals are measured by the same setup with the same scanning step. Obviously, the intensity response of the original DLA is very noisy and its peak is hard to define. On the contrary, the See-through DLA produces a much smoother signal, which proves the concept of increasing the zero-order diffraction efficiency to reduce the disturbance. The FWHM of the axial response of the See-through DLA is $26.38\ \mu\text{m}$, which is equivalent to the confocal plane response of an objective with an NA of 0.16 according to Equation 2.135. There is a slight decrease in the FWHM of the axial response compared to the confocal response of a 0.15 NA objective, which also fits well with the simulation result by FFT-RSI in Figure 2.9.

Subsequently, a multi-spot confocal measurement can be performed by raster scanning of the sample in the lateral direction. Figure 4.19 shows such a measurement of a coin by the See-through DLA and the 5×0.15 NA objective. The DLA produces an 11×11 spot array with a pitch of $100\ \mu\text{m}$. The total measurement area is $1.1\ \text{mm}\times 1.1\ \text{mm}$. The coin is moved by the piezo stage with a lateral scanning step of $3\ \mu\text{m}$ and an axial scanning step of $1.2\ \mu\text{m}$. The axial peak at each lateral position is fitted by a quadratic function and its center

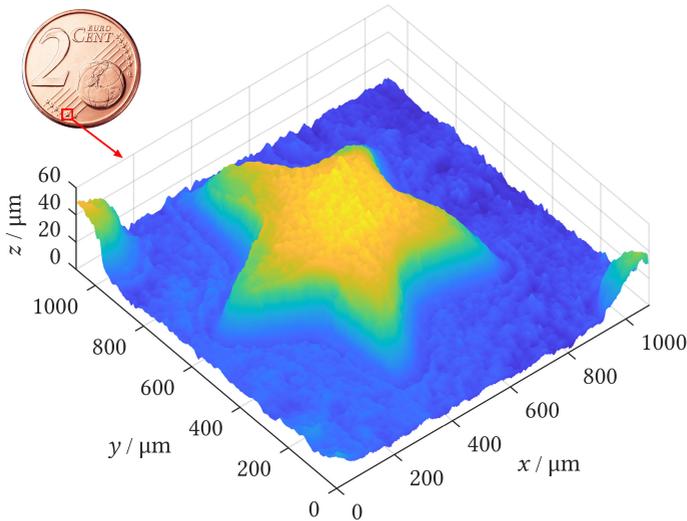


Figure 4.19: Confocal surface measurement of a star on a 2 cent coin by the 5×0.15 NA objective with the See-through DLA.

is used as the altitude at this position. The height map is filtered by a moving average filter with a window size of 5 pixels to eliminate the artifacts where the reflected signal is low, for example, at steep slopes.

4.3.2 Axial response of Direct-imaging DLAs

The test setup for the axial response of the Direct-imaging DLAs is also the same as Figure 4.15. The central pixel of each spot is used as the pinhole, and its intensity is recorded during the scanning.

Similarly, a mirror is placed underneath the setup as the sample. It is moved axially to measure the confocal axial responses of the Direct-imaging DLA. The measurement result of the central spot in the DLA which produces a 5×5 array with an NA of 0.7 and a pitch of $100\ \mu\text{m}$ is shown in Figure 4.20 (a). The FWHM of the confocal axial response is $2.24\ \mu\text{m}$. Compared to the axial response of the See-through DLA in Figure 4.18 (b), the FWHM is significantly reduced although both signals are measured by the same 5×0.15 NA objective,

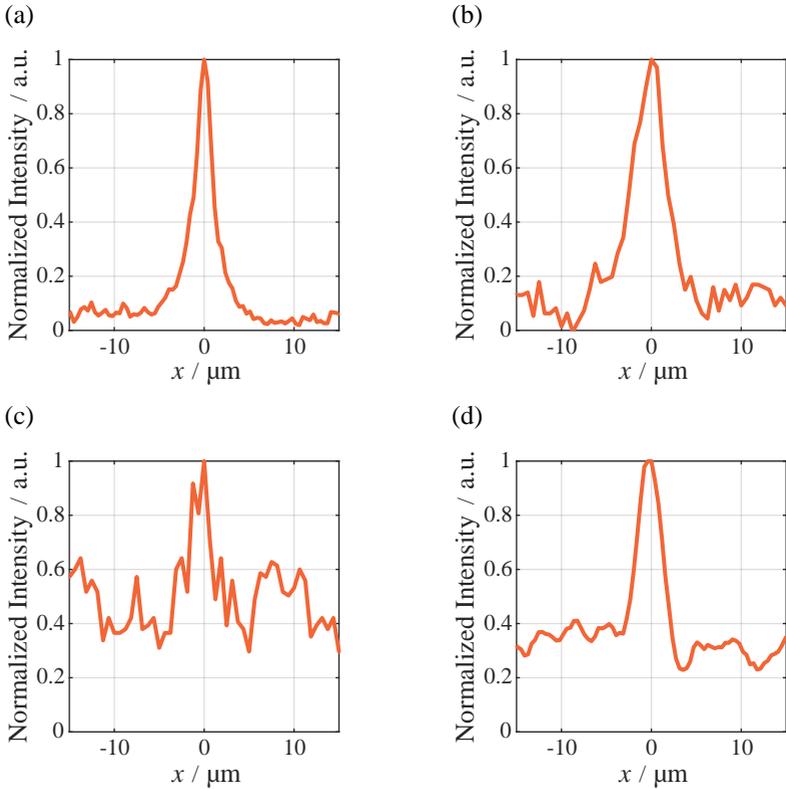


Figure 4.20: Confocal axial response of the Direct-imaging DLA and the 5×0.15 NA objective when a mirror is measured. (a) Central spot in the 5×5 spot array with $100 \mu\text{m}$ pitch. (b) Central spot in the 7×7 spot array with $100 \mu\text{m}$ pitch. (c) Central spot in the 9×9 spot array with $100 \mu\text{m}$ pitch. (d) Central spot in the 25×25 spot array with $200 \mu\text{m}$ pitch.

which proves that the Direct-imaging DLA acts exactly as a finite-conjugate lens array and they can perform high-NA confocal surface measurement.

However, with the increase of the number of spots, the noise of the signal will also increase. Figures 4.20 (b) and (c) show the confocal axial responses by a 7×7 and a 9×9 array respectively. It is obvious that the peaks suffer strong disturbance caused by cross-talks among the lens elements, which is also

predicted by the simulation in Figure 3.12. Changing the spot arrangement, for example, into a line, increasing the pitch of the array, or decreasing the NAs of the DLAs can reduce the cross-talk. For example, the axial response of the central spot in a 25×25 spot array with 0.5 NA and $200 \mu\text{m}$ pitch is measured as Figure 4.20 (d) shows. It appears that with a larger pitch the spot can still be distinguished in such a large grid.

4.3.3 Interference measurement by See-through DLAs

As discussed in the previous sections, the axial measurement sensitivity of the See-through DLAs is still limited by the objective. To overcome such a limitation and reduce the axial measurement uncertainties, the idea of interference is utilized. As one of the most precise distance measurement techniques, interference can improve the axial measurement sensitivity of the previous confocal microscope setup based on the See-through DLAs.

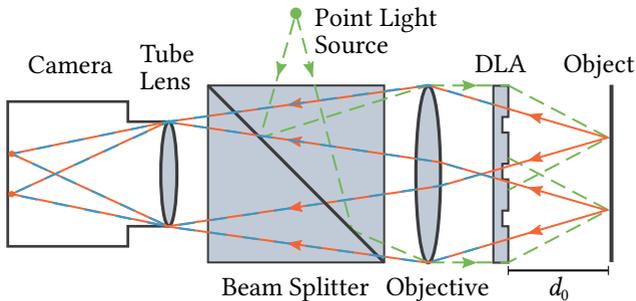


Figure 4.21: Setup of the confocal interference microscope based on the See-through DLA.

The concept is shown in Figure 4.21. Similarly, the DLA focuses the plane-wave illumination into a spot array. The spots are reflected by the sample as the probe beam, which is denoted as the red lines in the figure. Meanwhile, due to the reflection on the surface of the phase mask, a reflected conjugate wave is formed, which is represented by the blue lines in the figure. The conjugate wave is used as the reference beam, which has the same wave-front as the probe beam when the object is in focus. As a result, when the probe beam and

the reference beam is superimposed with each other, interference occurs on the image plane. The interference fringes will overlap with the confocal signal, and it can increase the sensitivity for surface measurement. The reflectance of the DLA is determined by the material itself, which is around 4% for fused silica at an incident angle of 90° . The reflectance can be further controlled by coatings.

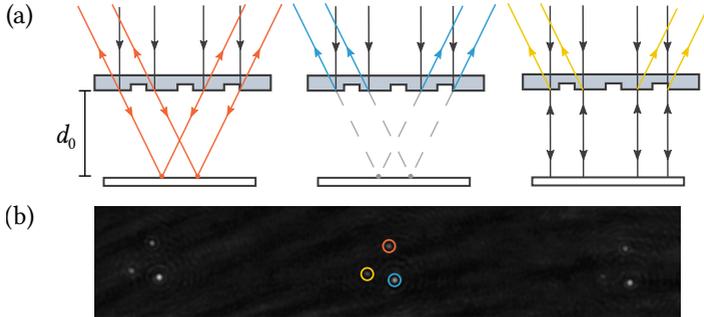


Figure 4.22: Reflection from the DLA and the sample surface. (a) Focusing wave, reflected conjugate wave and reflected wave by the plane surface. (b) Spots in the image formed by different waves when a piece of glass is placed as the sample.

When a surface is placed on the focal plane, the probe beam and the reference beam produce two sets of spots in the image, which are shown in Figure 4.22 as the red and blue lines respectively. The yellow spot is from the direct reflection by the plane surface, which has the same phase as the probe beam and will not overlap with spots when the sample has a rough surface or is placed with an angle. By mechanical alignment of the positions of the beam splitter, the light source and the DLA in the setup, the reference and probe spots in the image can superimpose with each other. Correspondingly, the reference and probe beams overlap with each other perfectly. When the coherence length of the laser is long enough, they will interfere with each other with an optical path difference of twice the working distance, which equals $2d_0$ when the sample is in focus. Therefore, the phase shift for the interference between the probe beam and the reference beam is $4\pi d_0/\lambda$.

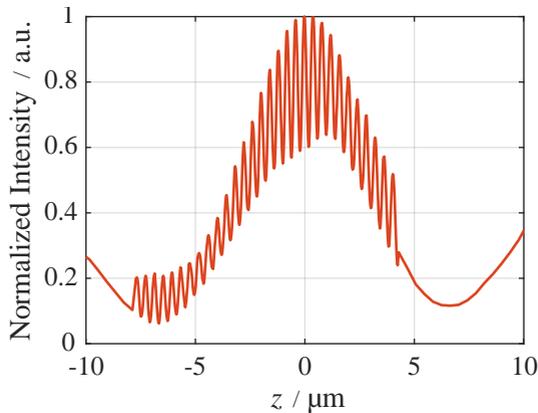


Figure 4.23: Confocal interference axial response by the See-through DLA and the 10×0.3 NA objective.

Figure 4.23 shows a typical confocal interference axial response when a mirror is measured at a wavelength of 785 nm with the single-frequency laser as the light source. Interference fringes with oscillations can be observed in the figure. The fringes disappear on the edge because the axial scanning step is set to a larger value to save the measurement time. The confocal peak has large side lobes due to the aberrations when the sample is imaged through the phase mask by the objective. Nevertheless, the confocal peak can help to solve the phase ambiguity of the interference fringes. The highest peak of the signal is fitted by a polynomial and the height is calculated. In this way, the range for determination of the axial position is reduced to half of the wavelength, instead of the broad confocal peak by the objective. Therefore, the concept can help to increase the axial sensitivity of the confocal microscope based on the See-through DLAs for surface measurement.

4.3.4 Measurement of a step height target

To evaluate the performance of various DLA concepts for axial surface measurement, a calibrated step height target (VLSI SHS-9400QC) is used as the

sample. The target is a piece of quartz coated with chromium with a precisely etched convex bar on its plane surface. The bar has a certified height of 925.5 ± 5.4 nm which is determined by the two-sigma rule. The step height target and its cross section is shown in Figure 4.24. The target is scanned axially by a piezo stage and confocal peaks of the spots are recorded to determine the heights at each position. The height of the bar is measured by following the procedures [VLS10, Gro17].

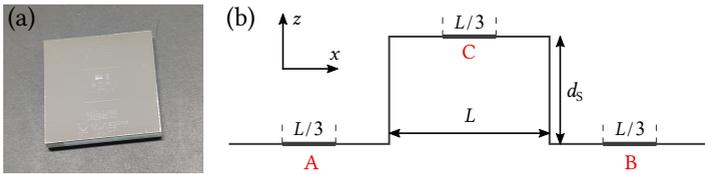


Figure 4.24: The step height standard calibration target. (a) Picture of the target. (b) Cross section of the measurement area.

Along the cross section of the bar, three sections of lines are measured, which are noted as A, B and C in Figure 4.24 (b). Each section has a length of one thirds of the width of the bar $L = 100 \mu\text{m}$. The heights of the lines are measured by discrete sampling points with an interval of $3 \mu\text{m}$ in the experiment. At each measurement point, the confocal signal is fitted by a polynomial. The peak of the polynomial is used as the height at this point. Then, the measured heights in the three sections are fitted by the following equation

$$z = \begin{cases} a_0x + a_1, & x \in (A, B), \\ a_0x + a_2, & x \in (C), \end{cases} \quad (4.1)$$

where z is the measured height, x is the lateral position and a_0 , a_1 and a_2 are the line fitting parameters. So the height of the bar can be easily calculated by

$$d_s = \frac{a_2 - a_1}{\sqrt{a_0^2 + 1}}. \quad (4.2)$$

Afterwards, the process is repeated 10 times along the y direction with an interval of $10\ \mu\text{m}$. The mean value and standard deviation of the measured heights are used for comparison with the ground-truth value. The measurement is carried out with different DLA concepts and also a commercial system (Leica DCM 3D) as a reference. The results are shown in Table 4.3.

Table 4.3: Measurement of the step height target.

Device	Method	$d_{S,\text{mean}} / \text{nm}$	σ_S / nm
Ground truth	Phase-shift interferometry	925.5	2.7
Commercial System	Confocal $5\times 0.15\ \text{NA}$	983.4	44.1
	Confocal $50\times 0.5\ \text{NA}$	981.5	11.1
See-through DLA	Confocal $5\times 0.15\ \text{NA}$	749.3	275.7
	Confocal $10\times 0.3\ \text{NA}$	960.6	49.2
	Interference $10\times 0.3\ \text{NA}$	904.7	9.7
Direct-imaging DLA	Confocal $5\times 0.15\ \text{NA}$	917.5	49.9

The major source of the measurement uncertainty for the DLA-based confocal microscopes is the stray light in the systems. The stray light mainly comes from other diffraction orders of the DLAs. For multi-level phase masks, the first-order diffraction efficiency is determined by the number of phase levels on the mask, which is given by the following equation [Swa89],

$$\eta_m^N = \left[\frac{\sin(m-1)\pi}{(m-1)\pi} \right]^2 \left[\frac{\sin(\pi/N)}{\pi/N} \right]^2, \quad (4.3)$$

where m is the diffraction order, N is the number of phase levels and η_m^N is the diffraction efficiency for the specific order at the designed wavelength. Therefore, as calculated by the above equation, a binary phase mask has a first-order diffraction efficiency of 41% at maximum in theory. Besides, the -1 diffraction order, which is the divergent wave in this case, always has the same energy as the $+1$ order. All the unwanted diffraction orders become stray light and eventually interfere with the confocal signals due to the coherence of the laser. The diffraction efficiency can be significantly improved by increase

of the phase levels. For example, a four-step phase mask has a first-order diffraction efficiency of 81% according to Equation 4.3.

The other part of the stray light comes from the reflection from the lens elements of the objective, the DLA glass surfaces and other optical components in the system. Because the objective is not designed for collimation from the back side, there is considerable reflection from the rear lens elements of the objective. Besides, the phase mask is not coated which has a reflectance of around 4% according to the Fresnel's equations [Fre23]. Using a specifically designed coating and polarization optics can help to alleviate the problem.

For the See-through DLA, the measurement uncertainty is directly influenced by the objective. The 10×0.3 NA objective performs much better than the 5×0.15 NA objective because the former one provides not only a smaller FWHM of the axial response, but also a higher SNR of the confocal signal. Although the See-through DLA mitigates the disturbance in reflected-light microscopes, there is still residual disturbance. Besides, the sample has to be imaged through the phase mask by the objective. This also introduces extra aberrations for the objective which affect the confocal axial response.

For the Direct-imaging DLA, the NA of the objective does not have much influence on the quality of the confocal signal, since only the low-NA spots in the intermediate image needs to be collected. However, disturbance and stray light become more significant because the signal intensity in the intermediate image is rather low due to the limited diffraction efficiency. Therefore, although its FWHM of the axial confocal response is smaller than that of the See-through DLA, the measurement result has not been remarkably better.

Last but not least, the interference concept of the See-through DLA does increase the sensitivity and reduce the uncertainty in the measurement, although it still suffers from the stray light. The overlapping with the stray light adds a random phase on the signal and aggravates the ambiguity problem of the interference fringes. In the experiment, the maxima of the confocal signals at each line sections in Figure 4.24 (b) are averaged. Then, the interference fringe closest to the average maximum is used to determine the height at each measurement point. Nevertheless, the proposed measurement setups

and concepts based on the DLAs approach the performance of the commercial 3D confocal metrology system. Moreover, they can be further upgraded by better engineering with the previously mentioned improvements, such as the multi-level phase mask and the matched coatings.

4.4 Fluorescence measurement

In this section, the application of the See-through DLAs in fluorescence microscopy is presented. Compared to their application in surface measurement, the sample is composed of point-like objects. Therefore, the resolution in both the lateral and axial directions can be enhanced by the See-through DLAs according to Equation 2.128 and 2.129. Besides, due to the emission filter in a fluorescence microscope, which is used to filter out the excitation light, all the stray light in the above-mentioned setups is filtered out as well. Subsequently, the See-through DLAs can image the fluorescent samples with high resolution in all directions with excellent SNR. The experiment setup of the confocal fluorescence microscope based on the See-through DLAs is shown in Figure 4.25. It is very similar to the surface measurement setup in Figure 4.9 only with extra excitation and emission filters. Besides, the beam splitter is changed to a dichroic mirror in order to collect more emission light from the sample.

First, the See-through DLA and the original DLA without the plane-wave component are again compared in fluorescence intensity during the measurement. Due to the difficulty in finding fluorescent dyes with excitation wavelength at 785 nm, the DOE-785 is used at a wavelength 633 nm which produces degraded spots with a measured lateral FWHM of $1.24\ \mu\text{m}$ [Li21e]. Nevertheless, it can still reveal the differences in signal strength of the two DLAs during fluorescence measurement. The fluorescent beads (SPHERO 8-Peaks Rainbow Calibration Particles) with diameters from $3\ \mu\text{m}$ to $3.4\ \mu\text{m}$ are used as the sample. During the experiment, the same bead is focused by the spots produced by the two DLAs successively with an excitation laser wavelength of 633 nm (Thorlabs LP633-SF50) and an emission wavelength of 672 nm to 712 nm by the corresponding filters and dichroic mirror (Edmund Optics CY5 Fluorescence Filter Set). The measurement results are shown in

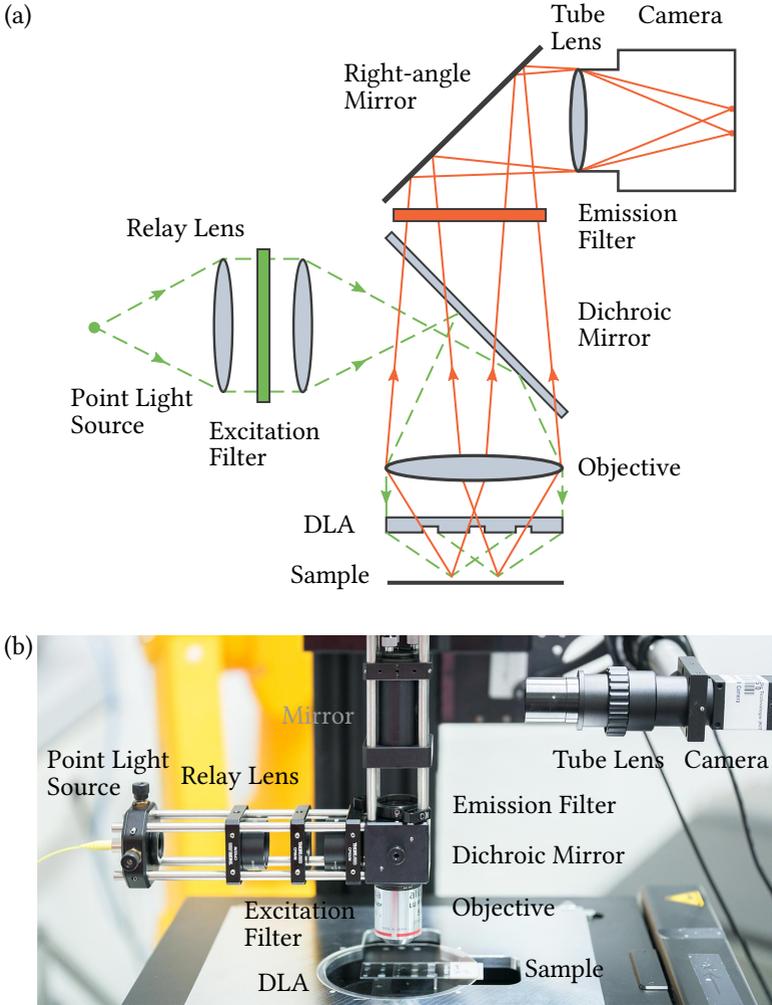


Figure 4.25: Measurement setup of a confocal fluorescence microscope based on the See-through DLAs. The sample is scanned laterally and an image is reconstructed. (a) A schematic of the experiment setup. (b) A picture of the experiment setup.

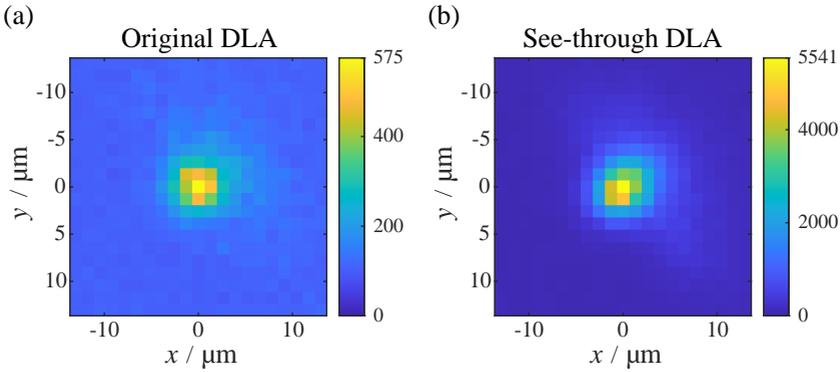


Figure 4.26: Comparison of the fluorescence intensity by focusing at the same fluorescent bead (SPHERO Rainbow Particles) with an excitation wavelength of 633 nm and an emission wavelength of 672 nm to 712 nm. (a) The original DLA with a peak gray value of 575. (b) The See-through DLA with a peak gray value of 5541.

Figure 4.26. By averaging the central 3×3 pixels, the mean gray value of the fluorescent signal with the original DLA is 450, while the one with the See-through DLA is 3946. The fluorescent signal of the See-through DLA is about 9 times brighter than that of the original DLA. Consequently, the resulting SNR of the confocal scanning image taken by the See-through DLA is much higher than that taken by the original DLA.

Next, the fluorescent beads are measured by the See-through DLA in the DOE-488 which produces a 41×41 spot array with a pitch of $75 \mu\text{m}$ and a working distance of $1001 \mu\text{m}$. The central spot produced by the DLA has a lateral FWHM of $0.303 \mu\text{m}$ and an axial FWHM of $0.866 \mu\text{m}$, which also corresponds to an NA of 0.83 [Li22]. The excitation wavelength is 488 nm and the emission wavelength is 510 nm to 550 nm with the corresponding filters (Thorlabs MDF-FITC Filter Set). The beads are dried out on a slide and they are imaged through air by the 5×0.15 NA objective and a 100 mm tube lens. The camera has a pixel size of $4.5 \mu\text{m}$. A window of 3×3 pixels around the center of each spot in the image is used as the pinhole. The measurement result is shown in Figure 4.27 (a). The total measurement area is $2.5 \text{ mm} \times 2.5 \text{ mm}$ due to the limitation of the camera sensor size. The scanning step is $0.3 \mu\text{m}$. In the confocal image, the round shape of the beads can be easily seen, and individual

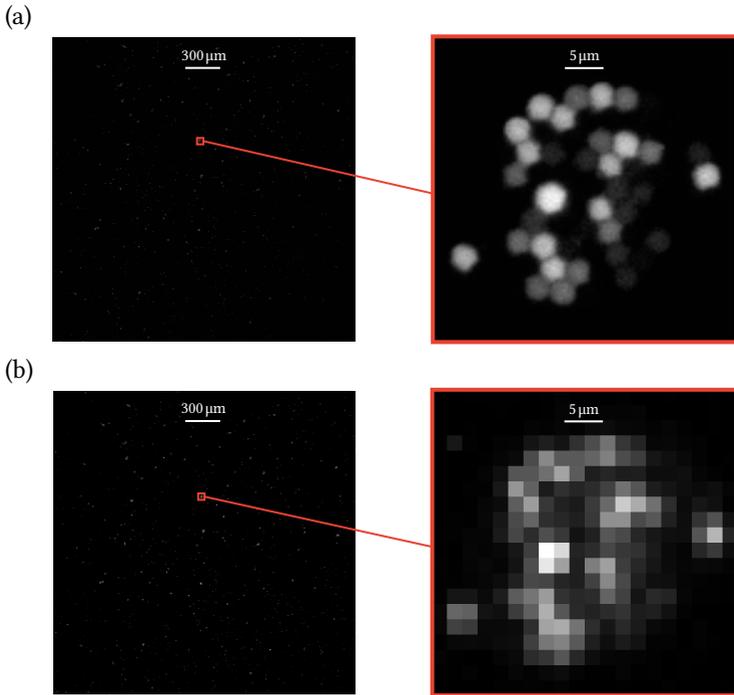


Figure 4.27: Measurement of the fluorescent beads (SPHERO Rainbow Particles) with diameters around $3\ \mu\text{m}$ by the $5\times 0.15\ \text{NA}$ objective. (a) Confocal scanning image by the See-through DLA with a scanning step of $0.3\ \mu\text{m}$. (b) Wide-field image by the objective solely.

beads with different intensities can be clearly distinguished. Compared to the wide-field image taken by the same objective in 4.27 (b), it is obvious that the resolution has been significantly enhanced by the DLA.

The sample can also be scanned axially and a 3D confocal image can be reconstructed. The volumetric plot is shown in Figure 4.28 with a lateral scanning step of $0.3\ \mu\text{m}$ and an axial scanning step of $1\ \mu\text{m}$. The beads have ellipsoidal shapes due to the fact that the PSF has a larger FWHM in the axial direction.

In addition, another kind of fluorescent beads (FluoSpheres Sulfate Microspheres 505/515) with diameters of $20\ \text{nm}$ is measured. The measurement

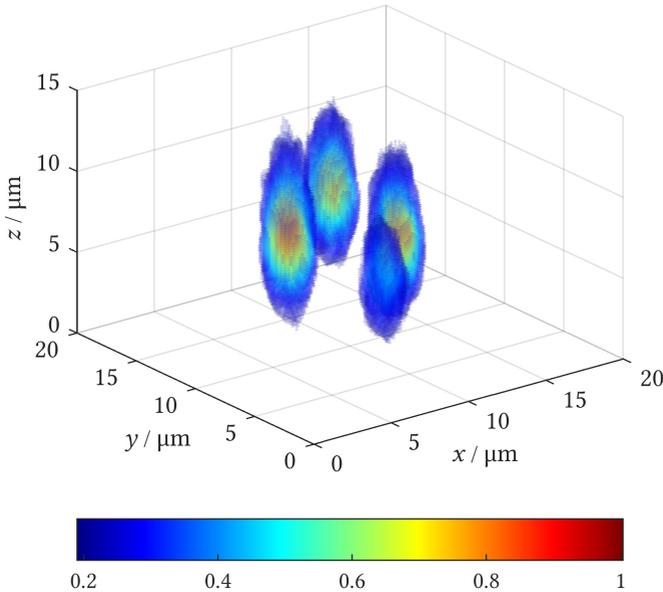


Figure 4.28: 3D confocal image of the fluorescent beads (SPHERO Rainbow Particles) by the See-through DLA and the 5×0.15 NA objective. The lateral scanning step is $0.3 \mu\text{m}$ and the axial scanning step is $1 \mu\text{m}$.

result by the same 5×0.15 NA objective and the 100 mm tube lens is shown in Figure 4.29. The total measurement area is $2.5 \text{ mm} \times 2.5 \text{ mm}$ with a scanning step of $0.2 \mu\text{m}$. The beads are also dried out on a slide. Due to the small sizes, individual beads cannot be distinguished. Instead, they accumulate and form a layer of film with small cracks when they are dried. Such fine structures can be clearly resolved by the confocal scanning with the DLA in Figure 4.29 (a), while they are invisible in the wide-field image in Figure 4.29 (b).

Furthermore, the objective is changed to the one with a smaller NA of 0.07. The same sample is measured as Figure 4.30 shows. The measurement area is $3 \text{ mm} \times 3 \text{ mm}$ with a scanning step of $0.3 \mu\text{m}$. According to Equation 1.3 and the Rayleigh criterion, this results in an SBP of around 300 megapixels, which is much higher than the SBPs of common commercial microscopes with several megapixels [Bia17, Zhe16]. By comparison of the same measurement

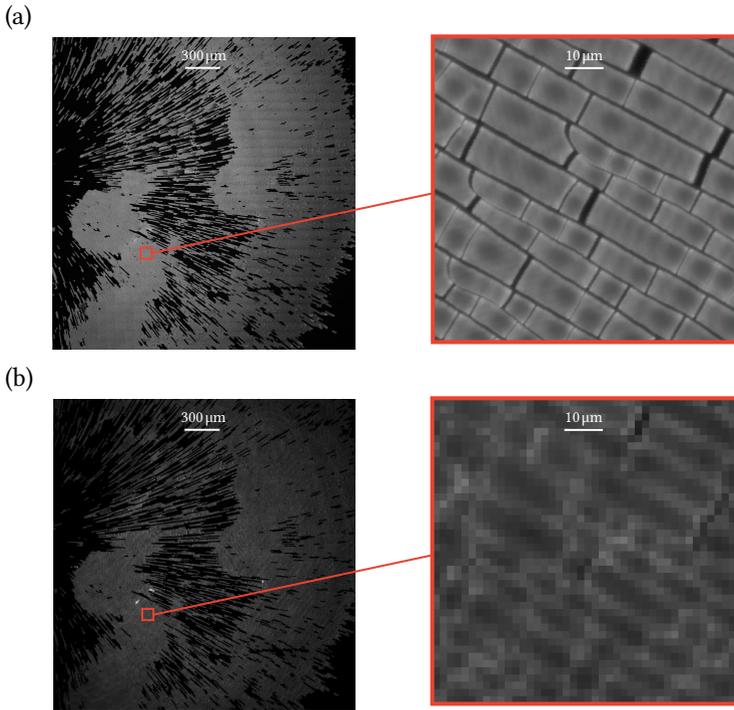


Figure 4.29: Measurement of the fluorescent beads (FluoSpheres Sulfate Microspheres 505/515) by the 5×0.15 NA objective. The total measurement area is $2.5\text{ mm} \times 2.5\text{ mm}$. (a) Confocal scanning image by the See-through DLA with a scanning step of $0.2\text{ }\mu\text{m}$. (b) Wide-field image by the objective solely.

region on the right side of Figure 4.30 (a) with that of Figure 4.29 (a), there is no significant reduction in the resolution although the NA of the objective is reduced by more than a half, while the wide-field image in Figure 4.30 (b) is completely blurred. However, there is a reduction in the contrast of the image, because less light is collected by the objective with a lower NA and the background remains unchanged. In this case, the contrast can be increased by using better filters and low-noise cameras. Besides, background subtraction in the post-processing can also help to improve the image contrast.

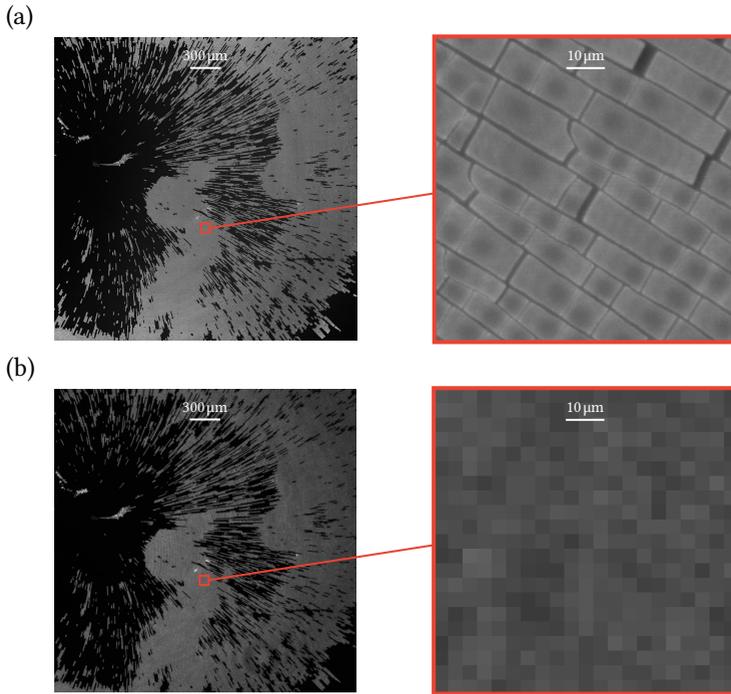


Figure 4.30: Measurement of the fluorescent beads (FluoSpheres Sulfate Microspheres 505/515) by the 2.5×0.07 NA objective. The total measurement area is $3\text{ mm} \times 3\text{ mm}$. (a) Confocal scanning image by the See-through DLA with a scanning step of $0.3\text{ }\mu\text{m}$. (b) Wide-field image by the objective solely.

Moreover, resolution of the proposed setup is not sensitive to lens aberrations. To prove that, a simple doublet lens with a focal length of 30 mm and a diameter of 25.4 mm (Thorlabs AC254-030-A-ML) is used as the objective. The schematic and the simulated MTF of the doublet by Zemax OpticStudio at 488 nm wavelength are shown in Figure 4.31 (b) and (c). The doublet has considerable aberrations throughout the field. The modulation contrast at $20\text{ lp}\cdot\text{mm}^{-1}$ is below 10%, corresponding to a line width or resolution of $25\text{ }\mu\text{m}$, which is much worse than most camera objectives available in the market. However, the confocal scanning image by the doublet and the DLA in Figure 4.31 (a) still clearly resolves the fine details of the sample. Although

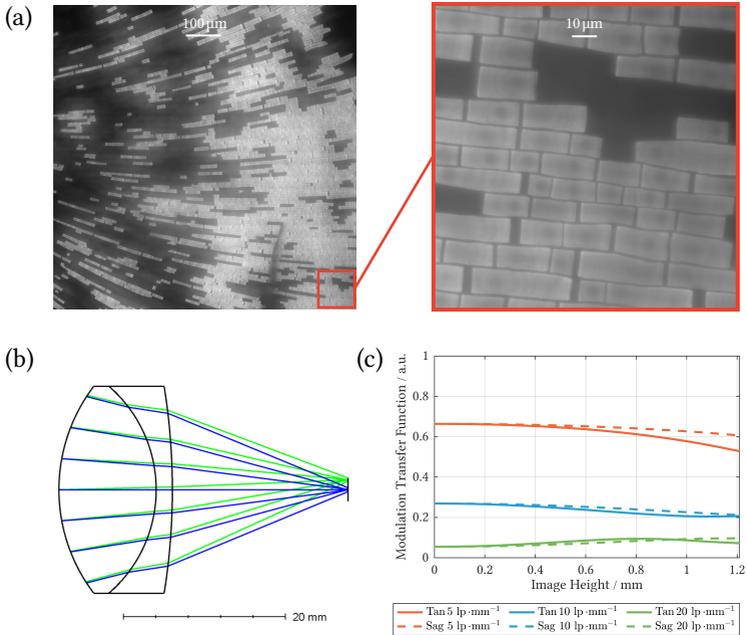


Figure 4.31: (a) Confocal scanning image of the fluorescent beads (FluoSpheres Sulfate Microspheres 505/515) by the See-through DLA and the doublet lens. The scanning step is $0.3 \mu\text{m}$ and the total measurement area is $1 \text{ mm} \times 1 \text{ mm}$. (b) Schematic of the doublet. (c) Simulated geometrical MTF of the doublet.

the background is brighter due to the aberrant rays, it can be alleviated by image dehazing algorithms in the post-processing.

In summary, high-resolution large-area fluorescence measurement has been demonstrated in this section. The See-through DLA has successfully measured a $3 \text{ mm} \times 3 \text{ mm}$ area by a 0.07 NA objective with the resolution equivalent to a 0.83 NA objective. Besides, resolution of the DLA-based confocal fluorescence microscope is not sensitive to the aberrations of the objective. Therefore, cheap objectives with large FOVs and moderate aberrations, e.g., full-frame camera lenses, can be used in the setup for high-resolution large-area imaging of the fluorescent samples, which can be potentially useful in large-scale biochemical assays.

5 Conclusions

The work in this thesis aims to extend the applications of DLAs in reflected-light confocal microscopy. In this closing chapter, the contributions and achievements of this thesis are summarized. An outlook is presented for possible improvements and further developments of this topic in the future.

5.1 Summary of the work

In traditional confocal microscopy, there is a trade-off between the spatial resolution and the FOV due to the limitations of the microscope objectives, which imposes restrictions on the SBPs of the systems. Commercial microscope objectives usually have SBPs in several megapixels, regardless their NAs and the magnifications [Bia17]. In previous research, DLAs are used to increase the FOV while maintaining the high resolution. They can produce high-NA spots over a large area, and thus the SBP of the measurement system can be significantly increased by scanning of the sample with the spots. However, previous DLAs have only measured semi-transparent samples in transmitted-light configurations, which are not favoured for surface metrology and fluorescence microscopy.

In this work, two different DLA concepts are proposed to overcome this limit and to enable the application of DLAs in reflected-light microscopes for measurement of opaque surfaces and fluorescent samples. The first See-through DLA design can reduce the field disturbances and let the projection spots be imaged through it. The second Direct-imaging DLA design emulates an array of finite-conjugate objectives which can provide high axial measurement sensitivity. A simulation framework based on the RSI is established, and the

idea of superposition is utilized to greatly reduce the number of variables and simplify the process of the optimization.

Two DLA prototypes are designed for wavelengths of 785 nm and 488 nm respectively. The binary phase masks are manufactured by electron-beam lithography. Different experiments are carried out to evaluate the DLAs. The See-through DLAs can produce spots with an NA up to 0.78 at a wavelength of 785 nm, and a lateral spatial frequency of $1024 \text{ lp}\cdot\text{mm}^{-1}$ has been measured by a 5×0.15 NA objective. Meanwhile, they have been shown to produce spots with an NA up to 0.83 at a wavelength of 488 nm, and a lateral spatial frequency of $1625 \text{ lp}\cdot\text{mm}^{-1}$ has been measured by the same 0.15 NA objective. By a 10×0.3 NA objective, the lateral spatial cut-off frequency can be increased to $2048 \text{ lp}\cdot\text{mm}^{-1}$, which so far demonstrates the highest spatial resolution measured by the DLAs to the best of the author's knowledge. For the Direct-imaging DLAs, the produced spots are measured with an NA up to 0.7 at a wavelength of 785 nm, and a lateral cut-off frequency of $1448 \text{ lp}\cdot\text{mm}^{-1}$ is measured by the 0.15 NA objective.

In the axial direction, a calibrated step height target with a nominal height of 925.5 nm is used as the reference to validate the performances of the DLA-based confocal setups in surface metrology. For the See-through DLA, the measured average height is 960.6 nm with a standard deviation of 49.2 nm by the 10×0.3 NA objective. Interference can be utilized to increase the measurement sensitivity of the See-through DLA, and the resulting measured average height is 904.7 nm with a standard deviation of 9.7 nm by the same objective. For the Direct-imaging DLA, the measured average height is 917.5 nm with a standard deviation of 49.9 nm by the 5×0.15 NA objective.

The application of the See-through DLAs in fluorescence microscopy has also been demonstrated. The fluorescent beads are measured by the DLA-based fluorescence microscope with a 2.5×0.07 NA objective at a wavelength of 488 nm. An area of $3 \text{ mm}\times 3 \text{ mm}$ is scanned by the spots produced by the DLA with an NA of 0.83. The corresponding SBP is about 300 megapixels, which is much higher than common commercial microscope systems with SBPs of several megapixels. The resolution of the proposed setup is also not sensitive

to lens aberrations. A high-resolution confocal image of the fluorescent beads has been shown by the DLA-based fluorescence microscope and a doublet lens.

In conclusion, this work has modified the DLAs in previous research and successfully applied them in reflected-light confocal microscopes. The new DLAs have shown the capabilities of high-resolution large-area measurement of opaque surfaces and fluorescent samples. They are potentially useful in both the biomedical inspection and industrial metrology.

5.2 Outlook

Although the experiments of the DLAs in this work have shown promising results, there are still several engineering problems which can be solved to improve the performance of the proposed setups.

As is discussed previously, the stray light has the most impact on the background noise of the confocal signals. Among the different sources of stray light, the unwanted diffraction orders of the DLAs have the most significant contribution. A multi-level phase mask can substantially increase the diffraction efficiency, and therefore decrease the amount of the stray light. As an example, a See-through DLA pattern with 3 phase levels is simulated as shown in Figure 5.1. The DLA generates a 21×21 spot array at a wavelength of 785 nm with a pitch of $20 \mu\text{m}$ and a feature size of $0.2 \mu\text{m}$, which is exactly the same as the binary See-through DLA in Figure 3.6. In the design process, instead of the binarization in Equation 3.4, the phase of the DLA is discretized by the following equation

$$\phi_S = \text{mod} \left(\left\lfloor \left(\frac{\arg(u_S)}{\pi} + B_N \right) \frac{N}{2} \right\rfloor, N \right) \frac{2\pi}{N}, \quad (5.1)$$

where N is the number of levels on the phase mask, and B_N is the corresponding discretization factor. After the same optimization process in Section 3.1.3, the following three-level See-through DLA with $N = 3$ has a working distance d_0 of $1099 \mu\text{m}$, a discretization factor B_3 of 0.01 and a plane-wave component weight W_S of 59. Similarly, by calculation of the central 5×5 unit cells in

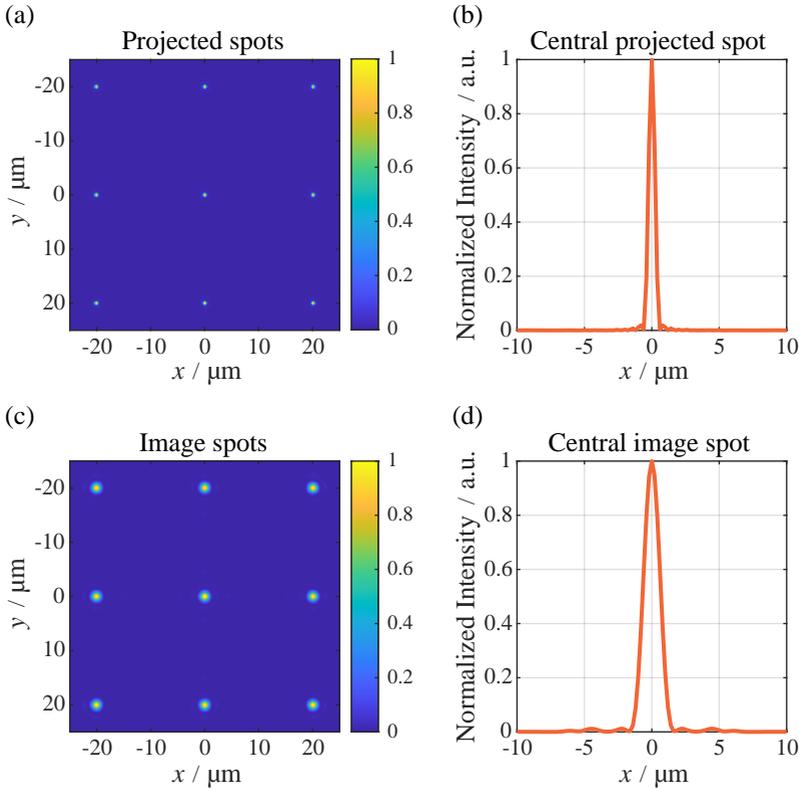


Figure 5.1: Simulation of spots generated and imaged through a See-through DLA with 3 phase levels. (a) The central 3×3 grid in the 21×21 spot array produced by the DLA on the object plane. (b) Cross section of the central spot on the object plane at $y = 0$. (c) The central 3×3 grid in the 21×21 spot array through the objective on the image plane. (d) Cross section of the central spot on the image plane at $y = 0$.

Figure 5.1 (c), the average PBR is 571. In comparison with the simulation of the binary pattern in Figure 3.6 (c) which has an PBR of 236, the PBR on the image plane has been significantly increased by the three-level phase mask and thus the image becomes much cleaner. Besides, by comparison of the peak intensities of the central projected spots in Figure 3.6 (b) and Figure 5.1 (b) with the same illumination power, the latter one is more than twice as high as the former one, although three-level DLA is designed to have a larger plane-wave

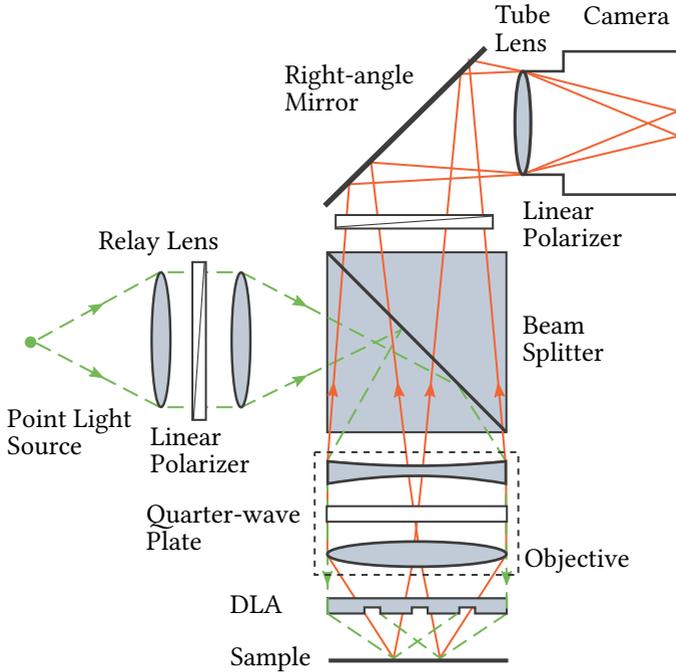


Figure 5.2: Polarization optics in the multi-spot confocal microscope based on the See-through DLAs to reduce the stray light.

component ratio W_S . This means that the energy is more concentrated due to the higher diffraction efficiency.

Therefore, it has been shown in the simulation that a multi-level phase mask can remarkably improve the signal quality of the DLA-based confocal microscope. However, the cost for a multi-level phase mask is considerably higher, since it requires multiple exposures and alignments in lithography. Nevertheless, the cost can still be reduced by mass production by compression molding.

Another source of stray light is the reflection from the surfaces of the optical components in the system. Such effect can be solved by specific coatings, customized objectives and polarization optics. For example, a quarter-wave

plate can be integrated into a customized objective in the DLA-based confocal microscope for surface measurement as shown in Figure 5.2. The linear polarizers are placed in an orthogonal configuration, so that only light with the polarization changed by the wave plate can enter the camera. In this way, all the reflections from the optical components above the quarter-wave plate can be eliminated. In addition, a coating can be applied on the upper surface of the phase mask to reduce the unwanted reflection.

Furthermore, the measurement speed of the current setup is generally limited by the raster scanning. The back and forth movements always have a speed limitation due to the mechanical acceleration and deceleration. This can be improved by deployment of a different scanning trajectory. Because the simulation framework based on the RSI in this work does not require periodicity, the spots projected by the DLA can be arranged in any pattern. Therefore, the diffractive lens elements can be placed in a different arrangement, for example, a Nipkow disk. Hence, the sample can be scanned by rotation of the diffractive lens disk to significantly increase the measurement speed.

In summary, the performance of the proposed DLA-based microscopes in this work can be further enhanced by the aforementioned improvements. Subsequently, the systems are promising for high-resolution large-area measurement at a low cost.

Bibliography

- [Abb73] ABBE, Ernst: “Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung”. In: *Archiv für mikroskopische Anatomie* 9.1 (1873), pp. 413–468.
- [Abr16] ABRAHAMSSON, Sara and GUSTAFSSON, Mats G. L.: “Multi-dimensional imaging using multi-focus microscopy”. U.S. pat. 9,477,091 B2. INSTITUTE, Howard Hughes Medical. Oct. 25, 2016.
- [Ada18] ADAM, Marie-Pierre; MÜLLENBROICH, Marie Caroline; GIOVANNA, Antonino Paolo Di; ALFIERI, Domenico; SILVESTRI, Ludovico; SACCONI, Leonardo and PAVONE, Francesco Saverio: “Confocal multispot microscope for fast and deep imaging in semicleared tissues”. In: *Journal of Biomedical Optics* 23.2 (2018), pp. 1–4. DOI: 10.1117/1.JBO.23.2.020503.
- [Arr10] ARRASMITH, Christopher L.; DICKENSHEETS, David L. and MAHADEVAN-JANSEN, Anita: “MEMS-based handheld confocal microscope for in-vivo skin imaging”. In: *Opt. Express* 18.4 (Feb. 2010), pp. 3805–3819. DOI: 10.1364/OE.18.003805.
- [Ben98] BENDICKSON, Jon M.; GLYTSIS, Elias N. and GAYLORD, Thomas K.: “Scalar integral diffraction methods: unification, accuracy, and comparison with a rigorous boundary element method with application to diffractive cylindrical lenses”. In: *Journal of the Optical Society of America A* 15.7 (July 1998), pp. 1822–1837. DOI: 10.1364/JOSAA.15.001822.
- [Bes18] BESSMELTSEV, VP; MAKSIMOV, MV; VILEIKO, VV; GOLOSHEVSKII, NV and TARENT’EV, VS: “Multichannel confocal microscope

based on a diffraction focusing multiplier with automatic synchronization of scanning”. In: *Optoelectronics, Instrumentation and Data Processing* 54.6 (2018), pp. 531–537.

- [Bet06] BETZIG, Eric; PATTERSON, George H; SOUGRAT, Rachid; LINDWASSER, O Wolf; OLENYCH, Scott; BONIFACINO, Juan S; DAVIDSON, Michael W; LIPPINCOTT-SCHWARTZ, Jennifer and HESS, Harald F: “Imaging intracellular fluorescent proteins at nanometer resolution”. In: *Science* 313.5793 (2006), pp. 1642–1645.
- [Bia17] BIAN, Liheng; SUO, Jinli; DAI, Qionghai and CHEN, Feng: “Fourier ptychography for high space-bandwidth product microscopy”. In: *Advanced Optical Technologies* 6.6 (2017), pp. 449–457. DOI: doi:10.1515/aot-2017-0054.
- [Boo50] BOOKER, H.G. and CLEMMOW, P.C.: “The concept of an angular spectrum of plane waves, and its relation to that of polar diagram and aperture distribution”. In: *Proceedings of the IEE - Part III: Radio and Communication Engineering* 97 (45 Jan. 1950), 11–17(6).
- [Bor13] BORN, Max and WOLF, Emil: *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light*. Elsevier, 2013.
- [Buc16] BUCHWALD, Jed Z and YEANG, Chen-Pang: “Kirchhoff’s theory for optical diffraction, its predecessor and subsequent development: the resilience of an inconsistent theory”. In: *Archive for History of Exact Sciences* 70.5 (2016), pp. 463–511. DOI: 10.1007/s00407-016-0176-1.
- [Cho13] CHOI, S.; KIM, P.; BOUTILIER, R.; KIM, M. Y.; LEE, Y. J. and LEE, H.: “Development of a high speed laser scanning confocal microscope with an acquisition rate up to 200 frames per second”. In: *Opt. Express* 21.20 (Oct. 2013), pp. 23611–23618. DOI: 10.1364/OE.21.023611.

- [Cle51] CLEMMOW, PC: “A method for the exact solution of a class of two-dimensional diffraction problems”. In: *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 205.1081 (1951), pp. 286–308. DOI: 10.1098/rspa.1951.0030.
- [Eng04] ENGELBERG, Yaakov M. and RUSCHIN, Shlomo: “Fast method for physical optics propagation of high-numerical-aperture beams”. In: *Journal of the Optical Society of America A* 21.11 (Nov. 2004), pp. 2135–2145. DOI: 10.1364/JOSAA.21.002135.
- [Fav92] FAVRO, Lawrence D.; THOMAS, Robert L.; KUO, Pao-Kuang and CHEN, Li: “Confocal Microscope”. U.S. pat. 5,162,941. UNIVERSITY, Wayne State. Nov. 10, 1992.
- [Fre19] FRESNEL, Augustin Jean: “Mémoire sur la diffraction de la lumière”. In: *Mémoires de l'Académie des sciences, tome V*. 1819, pp. 339–475.
- [Fre23] FRESNEL, Augustin Jean: “Mémoire sur la loi réflexion imprimée à la lumière polarisée”. In: *Académie des Sciences* 11 (1823), pp. 393–433.
- [FRS79] F.R.S., Lord Rayleigh: “XXXI. Investigations in optics, with special reference to the spectroscope”. In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 8.49 (1879), pp. 261–274. DOI: 10.1080/14786447908639684.
- [FRS97] F.R.S., Lord Rayleigh: “XXXVII. On the passage of waves through apertures in Plane screens, and allied problems”. In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 43.263 (1897), pp. 259–272. DOI: 10.1080/14786449708620990.
- [Goo05] GOODMAN, Joseph W: Introduction to Fourier optics. Roberts and Company Publishers, 2005.

- [Gro17] GROOT, Peter de and FITZGERALD, Danette: “Measurement, certification and use of step-height calibration specimens in optical metrology”. In: *Optical Measurement Systems for Industrial Inspection X*. Ed. by LEHMANN, Peter; OSTEN, Wolfgang and JR., Armando Albertazzi Gonçalves. Vol. 10329. International Society for Optics and Photonics. SPIE, 2017, pp. 328–336. DOI: 10.1117/12.2269800.
- [Hel94] HELL, Stefan W and WICHMANN, Jan: “Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy”. In: *Optics letters* 19.11 (1994), pp. 780–782.
- [Heu73] HEURTLEY, John C.: “Scalar Rayleigh–Sommerfeld and Kirchhoff diffraction integrals: A comparison of exact evaluations for axial points*”. In: *Journal of the Optical Society of America* 63.8 (Aug. 1973), pp. 1003–1008. DOI: 10.1364/JOSA.63.001003.
- [Hul12] HULSKEN, B.; VOSSEN, D. and STALLINGA, S.: “High NA diffractive array illuminators and application in a multi-spot scanning microscope”. In: *Journal of the European Optical Society-Rapid publications* 7 (2012).
- [Huy90] HUYGENS, Christiaan: *Traité de la lumière*. Leyden, 1690.
- [Ich96] ICHIHARA, Akira; TANAAMI, Takeo; ISOZAKI, Katsumi; SUGIYAMA, Yumiko; KOSUGI, Yasuhito; MIKURIYA, Kenta; ABE, Michio and UEMURA, Isao: “High-speed confocal fluorescence microscopy using a nipkow scanner with microlenses for 3-D imaging of single fluorescent molecule in real time”. In: *Bioimages* 4.2 (1996), pp. 57–62.
- [Ino02] INOUÉ, Shinya and INOUÉ, Ted: “Direct-view high-speed confocal scanner: the CSU-10”. In: *Methods in cell biology* 70 (2002), pp. 87–127. DOI: 10.1016/s0091-679x(02)70003-4.
- [Jah89] JAHNS, J; DOWNS, Maralene M; PRISE, ME; STREIBI, N and WALKER, Susan J: “Dammann gratings for laser beam shaping”. In: *Optical Engineering* 28.12 (1989), p. 281267.

- [Kim14] KIM, Yong-Hae; BYUN, Chun-Won; OH, Himchan; LEE, JaeWon; PI, Jae-Eun; KIM, Gi Heon; LEE, Myung-Lae; RYU, Hojun; CHU, Hye-Yong and HWANG, Chi-Sun: “Non-uniform sampling and wide range angular spectrum method”. In: *Journal of Optics* 16.12 (2014), p. 125710.
- [Kin96] KINO, Gordon S. and CORLE, Timothy R.: *Confocal scanning optical microscopy and related imaging systems*. Academic Press, 1996.
- [Kir83] KIRCHHOFF, Gustav: “Zur theorie der lichtstrahlen”. In: *Annalen der Physik* 254.4 (1883), pp. 663–695.
- [Koz15] KOZACKI, Tomasz and FALAGGIS, Konstantinos: “Angular spectrum-based wave-propagation method with compact space bandwidth for large propagation distances”. In: *Optics Letters* 40.14 (July 2015), pp. 3420–3423. DOI: 10.1364/OL.40.003420.
- [Lal68] LALOR, Éamon: “Conditions for the Validity of the Angular Spectrum of Plane Waves*”. In: *Journal of the Optical Society of America* 58.9 (Sept. 1968), pp. 1235–1237. DOI: 10.1364/JOSA.58.001235.
- [Lat92] LATIMER, Paul and CROUSE, Randy F.: “Talbot effect reinterpreted”. In: *Applied Optics* 31.1 (Jan. 1992), pp. 80–89. DOI: 10.1364/AO.31.000080.
- [Li19a] LI, Zheng: “Application of diffractive optical elements in confocal microscopy”. In: *Proceedings of the 2018 Joint Workshop of Fraunhofer IOSB and Institute for Anthropomatics, Vision and Fusion Laboratory*. Ed. by BEYERER, Jürgen and TAPHANEL, Miro. Vol. 40. KIT Scientific Publishing, 2019, pp. 25–46. DOI: 10.5445/KSP/1000094782.
- [Li19b] LI, Zheng; TAPHANEL, Miro; LÄNGLE, Thomas and BEYERER, Jürgen: “Application of DOE in confocal microscopy for surface measurement”. In: *IMEKO Joint TC1 - TC2 International Symposium on Photonics and Education in Measurement Science*. Ed. by

- ROSENBERGER, Maik; DITTRICH, Paul-Gerald and ZAGAR, Bernhard. Vol. 11144. International Society for Optics and Photonics. SPIE, 2019, pp. 254–261. DOI: 10.1117/12.2531610.
- [Li20a] LI, Zheng: “High-NA Confocal Measurement by Diffractive Optical Elements”. In: *Proceedings of the 2019 Joint Workshop of Fraunhofer IOSB and Institute for Anthropomatics, Vision and Fusion Laboratory*. Ed. by BEYERER, Jürgen and ZANDER, Tim. Vol. 45. KIT Scientific Publishing, 2020, pp. 85–94. DOI: 10.5445/KSP/1000118012.
- [Li20b] LI, Zheng and TAPHANEL, Miro: “Diffractive optical element, confocal microscope and method for designing a diffractive optical element”. Pat. WO 2020/099538. 2020.
- [Li20c] LI, Zheng; TAPHANEL, Miro; LÄNGLE, Thomas and BEYERER, Jürgen: “Direct-imaging DOEs for high-NA multi-spot confocal microscopy”. In: *tm - Technisches Messen* 87.s1 (2020), s40–s43. DOI: 10.1515/teme-2020-0017.
- [Li21a] LI, Zheng: “DOE-based Multi-spot Confocal Interference Microscope”. In: *Proceedings of the 2020 Joint Workshop of Fraunhofer IOSB and Institute for Anthropomatics, Vision and Fusion Laboratory*. Ed. by BEYERER, Jürgen and ZANDER, Tim. KIT Scientific Publishing, 2021. DOI: 10.5445/KSP/1000130397.
- [Li21b] LI, Zheng: “Mikroskopanordnung und Verfahren zum Messen einer Oberflächenstruktur einer Probe sowie diffraktives optisches Element”. Pat. DE102020205523. 2021.
- [Li21c] LI, Zheng; TAPHANEL, Miro; LÄNGLE, Thomas and BEYERER, Jürgen: “Direct-imaging DOEs for high-NA multi-spot confocal surface measurement”. In: *tm - Technisches Messen* 88.5 (2021), pp. 319–325. DOI: 10.1515/teme-2020-0103.
- [Li21d] LI, Zheng; TAPHANEL, Miro; LÄNGLE, Thomas and BEYERER, Jürgen: “High-resolution confocal microscopy with low-NA objectives based on diffractive lens arrays”. In: *Applied Optics* 60.22 (Aug. 2021), F1–F5. DOI: 10.1364/AO.423933.

- [Li21e] LI, Zheng; TAPHANEL, Miro; LÄNGLE, Thomas and BEYERER, Jürgen: “Resolution enhancement of low-NA objectives in confocal fluorescence microscopy by diffractive lens arrays”. In: *International Optical Design Conference 2021*. Ed. by CLARK, Peter P.; PFISTERER, Richard N.; REHN, Henning and THIBAUT, Simon. Vol. 12078. International Society for Optics and Photonics. SPIE, 2021, pp. 126–131. DOI: 10.1117/12.2603633.
- [Li22] LI, Zheng; TAPHANEL, Miro; LÄNGLE, Thomas and BEYERER, Jürgen: “Confocal fluorescence microscopy with high-NA diffractive lens arrays”. In: *Applied Optics* 61.3 (Jan. 2022), A37–A42. DOI: 10.1364/AO.442084.
- [Liu12] LIU, X.; STENAU, T. and BRENNER, K.-H.: “Diffractive micro lens arrays with overlapping apertures”. In: *Information Optics (WIO), 2012 11th Euro-American Workshop on*. IEEE, 2012, pp. 1–2.
- [Liu14] LIU, Xiyuan and BRENNER, Karl-Heinz: “High Resolution Wavefront Measurement with Phase Retrieval Using a Diffractive Overlapping Micro Lens Array”. In: *Fringe 2013*. Ed. by OSTEN, Wolfgang. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014, pp. 233–236.
- [Loh92] LOHMANN, Adolf W. and SINZINGER, Stefan: “Improved array illuminators”. In: *Applied Optics* 31.26 (Sept. 1992), pp. 5447–5452. DOI: 10.1364/AO.31.005447.
- [Loh96] LOHMANN, Adolf W.; DORSCH, Rainer G.; MENDLOVIC, David; ZALEVSKY, Zeev and FERREIRA, Carlos: “Space-bandwidth product of optical signals and systems”. In: *Journal of the Optical Society of America A* 13.3 (Mar. 1996), pp. 470–473.
- [Lun08] LUNDQUIST, Paul; ZHONG, Cheng-Frank; ZACCARIN, Denis; ZHAO, Pegian and TURNER, Stephen: “Methods and systems for analyzing fluorescent materials with reduced autofluorescence”. Pat. WO 2008/140758 A1. PACIFIC BIOSCIENCES OF CALIFORNIA, Inc. Nov. 20, 2008.

- [Mat09] MATSUSHIMA, Kyoji and SHIMOBABA, Tomoyoshi: “Band-Limited Angular Spectrum Method for Numerical Simulation of Free-Space Propagation in Far and Near Fields”. In: *Optics Express* 17.22 (Oct. 2009), pp. 19662–19673. DOI: 10.1364/OE.17.019662.
- [Mat21] MATSUMOTO, Naoya and OKAZAKI, Shigetoshi: “Image acquisition device and image acquisition method”. U.S. pat. 10,890,530 B2. HAMAMATSU UNIVERSITY SCHOOL OF MEDICINE NUC., Hamamatsu Photonics K.K. Jan. 12, 2021.
- [Max65] MAXWELL, J Clerk: “A dynamical theory of the electromagnetic field”. In: *Philosophical transactions of the Royal Society of London* 155 (1865), pp. 459–512.
- [Meh17] MEHRABKHANI, Soheil and SCHNEIDER, Thomas: “Is the Rayleigh-Sommerfeld diffraction always an exact reference for high speed diffraction algorithms?” In: *Optics Express* 25.24 (Nov. 2017), pp. 30229–30240. DOI: 10.1364/OE.25.030229.
- [Mel01] MELLIN, Stephen D. and NORDIN, Gregory P.: “Limits of scalar diffraction theory and an iterative angular spectrum algorithm for finite aperture diffractive optical element design”. In: *Optics Express* 8.13 (June 2001), pp. 705–722. DOI: 10.1364/OE.8.000705.
- [Min61] MINSKY, Marvin: “Microscopy apparatus”. U.S. pat. 3,013,467. Dec. 19, 1961.
- [Nip84] NIPKOW, Paul: “Elektrisches Teleskop”. German pat. 30105. Jan. 6, 1884.
- [Opp09] OPPENHEIM, Alan V. and SCHAFER, Ronald W.: *Discrete-Time Signal Processing*. 3rd. USA: Prentice Hall Press, 2009.
- [Ore14] OREOPOULOS, John; BERMAN, Richard and BROWNE, Mark: “Spinning-disk confocal microscopy: present technology and future trends”. In: *Quantitative Imaging in Cell Biology*. Ed. by WATERS, Jennifer C. and WITTMAN, Torsten. Vol. 123. *Methods in Cell Biology*. Academic Press, 2014, pp. 153–175. DOI: 10.1016/B978-0-12-420138-5.00009-4.

- [Ort12] ORTH, Antony and CROZIER, Kenneth: “Microscopy with microlens arrays: high throughput, high resolution and light-field imaging”. In: *Opt. Express* 20.12 (June 2012), pp. 13522–13531. DOI: 10.1364/OE.20.013522.
- [Ort14] ORTH, Antony and CROZIER, Kenneth B.: “High throughput multichannel fluorescence microscopy with microlens arrays”. In: *Opt. Express* 22.15 (July 2014), pp. 18101–18112. DOI: 10.1364/OE.22.018101.
- [Pac17] PACHECO, Shaun; WANG, Chengliang; CHAWLA, Monica K; NGUYEN, Minhkhloi; BAGGETT, Brend K; UTZINGER, Urs; BARNES, Carol A and LIANG, Rongguang: “High resolution, high speed, long working distance, large field of view confocal fluorescence microscope”. In: *Scientific reports* 7.1 (2017), pp. 1–10.
- [Pan12] PANG, Shuo; HAN, Chao; KATO, Mihoko; STERNBERG, Paul W. and YANG, Changhuei: “Wide and scalable field-of-view Talbot-grid-based fluorescence microscopy”. In: *Opt. Lett.* 37.23 (Dec. 2012), pp. 5018–5020. DOI: 10.1364/OL.37.005018.
- [Pan13] PANG, Shuo; HAN, Chao; ERATH, Jessej; RODRIGUEZ, Ana and YANG, Changhuei: “Wide field-of-view Talbot grid-based microscopy for multicolor fluorescence imaging”. In: *Opt. Express* 21.12 (June 2013), pp. 14555–14565. DOI: 10.1364/OE.21.014555.
- [Pet68] PETRÁŇ, Mojmír; HADRAVSKÝ, Milan; EGGER, M. David and GALAMBOS, Robert: “Tandem-Scanning Reflected-Light Microscope*”. In: *Journal of the Optical Society of America* 58 (May 1968), pp. 661–664. DOI: 10.1364/JOSA.58.000661.
- [Plo99] PLOEM, JOHAN S.: “Fluorescence microscopy”. In: *Fluorescent and Luminescent Probes for Biological Activity*. Ed. by MASON, W. T. London: Academic Press, 1999. Chap. 1, pp. 3–13.
- [Poi89] POINCARÉ, Henri and BLONDIN, Jules: *Leçons sur la théorie mathématique de la lumière: professées pendant le premier semestre 1887-1888*. Vol. 1. G. Carré, 1889, pp. 115–116.

- [Pom94] POMMET, Drew A.; MOHARAM, M. G. and GRANN, Eric B.: “Limits of scalar diffraction theory for diffractive phase elements”. In: *Journal of the Optical Society of America A* 11.6 (June 1994), pp. 1827–1834. DOI: 10.1364/JOSAA.11.001827.
- [Rit14] RITTER, André: “Modified shifted angular spectrum method for numerical propagation at reduced spatial sampling rates”. In: *Optics Express* 22.21 (Oct. 2014), pp. 26265–26276. DOI: 10.1364/OE.22.026265.
- [Rus09] RUSK, Nicole: “The fluorescence microscope”. In: *Nature Cell Biology* 11.1 (2009), S8–S9. DOI: 10.1038/ncb1941.
- [Sch20] SCHWEDT, Daniel: “Mikroskop und Verfahren zur Mikroskopie”. German pat. 102018127281 A1. GMBH, Carl Zeiss Microscopy. Apr. 30, 2020.
- [Sch96] SCHATZMAN, James C.: “Accuracy of the discrete Fourier transform and the fast Fourier transform”. In: *SIAM Journal on Scientific Computing* 17.5 (1996), pp. 1150–1166.
- [Sha49] SHANNON, C.E.: “Communication in the Presence of Noise”. In: *Proceedings of the IRE* 37.1 (1949), pp. 10–21. DOI: 10.1109/JRPROC.1949.232969.
- [She06] SHEN, Fabian and WANG, Anbo: “Fast-Fourier-transform based numerical integration method for the Rayleigh-Sommerfeld diffraction formula”. In: *Applied Optics* 45.6 (Feb. 2006), pp. 1102–1110. DOI: 10.1364/AO.45.001102.
- [She92] SHEPPARD, C. J. R. and HRYNEVYCH, M.: “Diffraction by a circular aperture: a generalization of Fresnel diffraction theory”. In: *Journal of the Optical Society of America A* 9.2 (Feb. 1992), pp. 274–281. DOI: 10.1364/JOSAA.9.000274.
- [Shi12] SHIMOBABA, Tomoyoshi; MATSUSHIMA, Kyoji; KAKUE, Takashi; MASUDA, Nobuyuki and ITO, Tomoyoshi: “Scaled angular spectrum method”. In: *Optics Letters* 37.19 (Oct. 2012), pp. 4128–4130. DOI: 10.1364/OL.37.004128.

- [Som04] SOMMERFELD, Arnold: *Mathematical Theory of Diffraction*. Boston, MA: Birkhäuser Boston, 2004, pp. 9–68. DOI: 10.1007/978-0-8176-8196-8_2.
- [Som54] SOMMERFELD, Arnold: *Optics: Lectures on Theoretical Physics, Volume IV*. Academic Press Inc., 1954.
- [Som96] SOMMERFELD, Arnold: “Mathematische theorie der diffraction”. In: *Mathematische Annalen* 47.2 (1896), pp. 317–374.
- [Ste16] STENAU, Tim and BRENNER, Karl-Heinz: “Diffractive Lenses with Overlapping Aperture A New Tool in Scanning Microscopy”. In: *Imaging Systems and Applications*. Optical Society of America. 2016, IT1F–1.
- [Ste17] STENAU, Tim: “Diffraktive Linsen mit überlappenden Aperturen in Simulation und Anwendung”. PhD thesis. Heidelberg University, Jan. 2017.
- [Sun16] SUN, Yangyang and PANG, Shuo: “Multi-perspective scanning microscope based on Talbot effect”. In: *Applied Physics Letters* 108.2 (2016), p. 021102.
- [Swa89] SWANSON, Gary J: *Binary optics technology: the theory and design of multi-level diffractive optical elements*. Tech. rep. Lincoln Laboratory, Massachusetts Institute of Technology, 1989.
- [Taf05] TAFLOVE, Allen and HAGNESS, Susan C: *Computational electrodynamics: the finite-difference time-domain method*. Artech house, 2005.
- [Tan02] TANAAMI, Takeo; OTSUKI, Shinya; TOMOSADA, Nobuhiro; KOSUGI, Yasuhito; SHIMIZU, Mizuho and ISHIDA, Hideyuki: “High-speed 1-frame/ms scanning confocal microscope with a microlens and Nipkow disks”. In: *Applied Optics* 41.22 (Aug. 2002), pp. 4704–4708. DOI: 10.1364/AO.41.004704.

- [Too06] TOOMRE, Derek and PAWLEY, James B.: “Disk-Scanning Confocal Microscopy”. In: ed. by PAWLEY, James B. *Handbook Of Biological Confocal Microscopy*. Boston, MA: Springer US, 2006, pp. 221–238. DOI: 10.1007/978-0-387-45524-2_10.
- [Vij15] VIJAYAKUMAR, Anand and BHATTACHARYA, Shanti: “Design of multifunctional diffractive optical elements”. In: *Optical Engineering* 54.2 (2015), pp. 1–7. DOI: 10.1117/1.OE.54.2.024104.
- [VLS10] VLSI STANDARDS: Application Note: Step Height Standards for use with KLA-Tencor Instruments. Rev.AB. 2010.
- [Wan19] WANG, Siran; LI, Zhiping; WU, Jianhua and WANG, Zhengpeng: “Accelerated near-field algorithm of sparse apertures by non-uniform fast Fourier transform”. In: *Optics Express* 27.14 (July 2019), pp. 19102–19118. DOI: 10.1364/OE.27.019102.
- [War19] WARREN, Craig; GIANNOPOULOS, Antonios; GRAY, Alan; GIANNAKIS, Iraklis; PATTERSON, Alan; WETTER, Laura and HAMRAH, Andre: “A CUDA-based GPU engine for gprMax: Open source FDTD electromagnetic simulation software”. In: *Computer Physics Communications* 237 (2019), pp. 208–218. DOI: 10.1016/j.cpc.2018.11.007.
- [Wey19] WEYL, Hermann: “Ausbreitung elektromagnetischer Wellen über einem ebenen Leiter”. In: *Annalen der Physik* 365.21 (1919), pp. 481–500.
- [Wil09] WILSON, Tony: “Confocal Microscopy”. In: *Biomedical Optical Imaging*. Ed. by FUJIMOTO, James G. and FARKAS, Daniel. Oxford: Oxford University Press, 2009. Chap. 1, pp. 3–28.
- [Wil11] WILSON, Tony: “Resolution and optical sectioning in the confocal microscope”. In: *Journal of microscopy* 244.2 (2011), pp. 113–121.
- [Wil84] WILSON, Tony and SHEPPARD, Colin: *Theory and practice of scanning optical microscopy*. Vol. 180. Academic Press London, 1984.

- [Wol64] WOLF, E. and MARCHAND, E. W.: “Comparison of the Kirchhoff and the Rayleigh-Sommerfeld Theories of Diffraction at an Aperture”. In: *Journal of the Optical Society of America* 54.5 (May 1964), pp. 587–594. DOI: 10.1364/JOSA.54.000587.
- [Wu10] WU, Jigang; CUI, Xiquan; ZHENG, Guoan; WANG, Ying Min; LEE, Lap Man and YANG, Changhui: “Wide field-of-view microscope based on holographic focus grid illumination”. In: *Opt. Lett.* 35.13 (July 2010), pp. 2188–2190. DOI: 10.1364/OL.35.002188.
- [Zha06] ZHANG, Fucui; PEDRINI, Giancarlo and OSTEN, Wolfgang: “Reconstruction algorithm for high-numerical-aperture holograms with diffraction-limited resolution”. In: *Optics Letters* 31.11 (June 2006), pp. 1633–1635. DOI: 10.1364/OL.31.001633.
- [Zha20a] ZHANG, Wenhui; ZHANG, Hao and JIN, Guofan: “Adaptive-sampling angular spectrum method with full utilization of space-bandwidth product”. In: *Optics Letters* 45.16 (Aug. 2020), pp. 4416–4419. DOI: 10.1364/OL.393111.
- [Zha20b] ZHANG, Wenhui; ZHANG, Hao; SHEPPARD, Colin J. R. and JIN, Guofan: “Analysis of numerical diffraction calculation methods: from the perspective of phase space optics and the sampling theorem”. In: *Journal of the Optical Society of America A* 37.11 (Nov. 2020), pp. 1748–1766. DOI: 10.1364/JOSAA.401908.
- [Zhe16] ZHENG, Guoan: *Fourier Ptychographic Imaging: A Matlab Tutorial*. Morgan & Claypool Publishers, 2016.
- [Zhu20] ZHU, Xufeng; FANG, Wei; LEI, Jian; LI, Zhangyin; XIE, Fei; CAO, Yaoyu; ZHANG, Yaping; QIN, Fei and LI, Xiangping: “Supercritical lens array in a centimeter scale patterned with maskless UV lithography”. In: *Optics Letters* 45.7 (Apr. 2020), pp. 1798–1801. DOI: 10.1364/OL.389702.

Publications

- [1] UMMETHALA, Sandeep; HARTER, Tobias; KÖHNLE, Kira; LI, Zheng; MÜHLBRANDT, Sascha et al.: “Wireless Transmission at 0.3 THz Using Direct THz-to-Optical Conversion at the Receiver”. In: *2018 European Conference on Optical Communication (ECOC)*. Sept. 2018, pp. 1–3. DOI: 10.1109/ECOC.2018.8535490.
- [2] FREUDE, Wolfgang; UMMETHALA, Sandeep; HARTER, Tobias; KÖHNLE, Kira; LI, Zheng et al.: “Wireless THz-to-optical conversion with an electro-optic plasmonic modulator”. In: *Light Conference 2019*. Light: Science & Applications, Light Conference (Chángchūn, China, July 16–18, 2019). 2019.
- [3] LI, Zheng: “Application of diffractive optical elements in confocal microscopy”. In: *Proceedings of the 2018 Joint Workshop of Fraunhofer IOSB and Institute for Anthropomatics, Vision and Fusion Laboratory*. Ed. by BEYERER, Jürgen and TAPHANEL, Miro. Vol. 40. KIT Scientific Publishing, 2019, pp. 25–46. DOI: 10.5445/KSP/1000094782.
- [4] LI, Zheng; TAPHANEL, Miro; LÄNGLE, Thomas and BEYERER, Jürgen: “Application of DOE in confocal microscopy for surface measurement”. In: *IMEKO Joint TC1 - TC2 International Symposium on Photonics and Education in Measurement Science*. Ed. by ROSENBERGER, Maik; DITTRICH, Paul-Gerald and ZAGAR, Bernhard. Vol. 11144. International Society for Optics and Photonics. SPIE, 2019, pp. 254–261. DOI: 10.1117/12.2531610.
- [5] NEGARA, Christian; LI, Zheng; LÄNGLE, Thomas and BEYERER, Jürgen: “Simplified Stokes polarimeter based on division-of-amplitude”. In: *IMEKO Joint TC1 - TC2 International Symposium on Photonics and Education in Measurement Science*. Ed. by ROSENBERGER, Maik;

- DITTRICH, Paul-Gerald and ZAGAR, Bernhard. Vol. 11144. International Society for Optics and Photonics. SPIE, 2019, pp. 316–325. DOI: 10.1117/12.2532399.
- [6] UMMETHALA, Sandeep; HARTER, Tobias; KÖHNLE, Kira; LI, Zheng; MÜHLBRANDT, Sascha et al.: “THz-to-optical conversion in wireless communications using an ultra-broadband plasmonic modulator”. In: *Nature photonics* 13.8 (2019), pp. 519–524. DOI: 10.1038/s41566-019-0475-6.
- [7] LI, Zheng: “High-NA Confocal Measurement by Diffractive Optical Elements”. In: *Proceedings of the 2019 Joint Workshop of Fraunhofer IOSB and Institute for Anthropomatics, Vision and Fusion Laboratory*. Ed. by BEYERER, Jürgen and ZANDER, Tim. Vol. 45. KIT Scientific Publishing, 2020, pp. 85–94. DOI: 10.5445/KSP/1000118012.
- [8] LI, Zheng and TAPHANEL, Miro: “Diffractive optical element, confocal microscope and method for designing a diffractive optical element”. Pat. WO 2020/099538. 2020.
- [9] LI, Zheng; TAPHANEL, Miro; LÄNGLE, Thomas and BEYERER, Jürgen: “Direct-imaging DOEs for high-NA multi-spot confocal microscopy”. In: *tm - Technisches Messen* 87.s1 (2020), s40–s43. DOI: 10.1515/teme-2020-0017.
- [10] LI, Zheng: “DOE-based Multi-spot Confocal Interference Microscope”. In: *Proceedings of the 2020 Joint Workshop of Fraunhofer IOSB and Institute for Anthropomatics, Vision and Fusion Laboratory*. Ed. by BEYERER, Jürgen and ZANDER, Tim. KIT Scientific Publishing, 2021. DOI: 10.5445/KSP/1000130397.
- [11] LI, Zheng: “Mikroskopianordnung und Verfahren zum Messen einer Oberflächenstruktur einer Probe sowie diffraktives optisches Element”. Pat. DE102020205523. 2021.
- [12] LI, Zheng; TAPHANEL, Miro; LÄNGLE, Thomas and BEYERER, Jürgen: “Direct-imaging DOEs for high-NA multi-spot confocal surface measurement”. In: *tm - Technisches Messen* 88.5 (2021), pp. 319–325. DOI: 10.1515/teme-2020-0103.

- [13] LI, Zheng; TAPHANEL, Miro; LÄNGLE, Thomas and BEYERER, Jürgen: “High-resolution confocal microscopy with low-NA objectives based on diffractive lens arrays”. In: *Applied Optics* 60.22 (Aug. 2021), F1–F5. DOI: 10.1364/AO.423933.
- [14] LI, Zheng; TAPHANEL, Miro; LÄNGLE, Thomas and BEYERER, Jürgen: “Resolution enhancement of low-NA objectives in confocal fluorescence microscopy by diffractive lens arrays”. In: *International Optical Design Conference 2021*. Ed. by CLARK, Peter P.; PFISTERER, Richard N.; REHN, Henning and THIBAUT, Simon. Vol. 12078. International Society for Optics and Photonics. SPIE, 2021, pp. 126–131. DOI: 10.1117/12.2603633.
- [15] LI, Zheng; TAPHANEL, Miro; LÄNGLE, Thomas and BEYERER, Jürgen: “Confocal fluorescence microscopy with high-NA diffractive lens arrays”. In: *Applied Optics* 61.3 (Jan. 2022), A37–A42. DOI: 10.1364/AO.442084.

A Kirchhoff's Diffraction Formulation

Inspired by the work of Helmholtz on acoustic waves, Kirchhoff deployed Green's theorem to tackle the problem [Buc16]. The starting point is the well-known Green's second identity, which is denoted as

$$\iiint_V P \nabla^2 Q - Q \nabla^2 P \, dV = \iint_S P \frac{\partial Q}{\partial \mathbf{n}} - Q \frac{\partial P}{\partial \mathbf{n}} \, dS, \quad (\text{A.1})$$

where P and Q are functions with continuous first and second derivatives, V is the integration volume surrounded by the surface S , and $\partial/\partial \mathbf{n}$ is the derivative along the outward normal direction to the surface S .

Next, P is chosen to be the scalar field $u(\mathbf{r})$ in Equation 2.23, and Q to be the Green's function of the same equation, which satisfies

$$\nabla^2 u(\mathbf{r}) + k^2 u(\mathbf{r}) = 0, \quad (\text{A.2})$$

$$\nabla^2 G(\mathbf{r}, \mathbf{r}') + k^2 G(\mathbf{r}, \mathbf{r}') = \delta(\mathbf{r} - \mathbf{r}'), \quad (\text{A.3})$$

where $\delta(\mathbf{r} - \mathbf{r}')$ is the Dirac delta function. The Green's function $G(\mathbf{r}, \mathbf{r}')$ of the Helmholtz equation is known to be

$$G(\mathbf{r}, \mathbf{r}') = \frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r} - \mathbf{r}'|}. \quad (\text{A.4})$$

Note that $G(\mathbf{r}, \mathbf{r}')$ is everywhere non-singular except at the point where $\mathbf{r}' = \mathbf{r}$. Thus, in order to apply it to the Green's second identity in Equation A.1, a special volume as shown in Figure A.1 has to be constructed to exclude the point at \mathbf{r} . The volume V has excluded an infinitesimal sphere surrounding the point \mathbf{r} with a radius of ϵ . With such configuration, surface surrounding

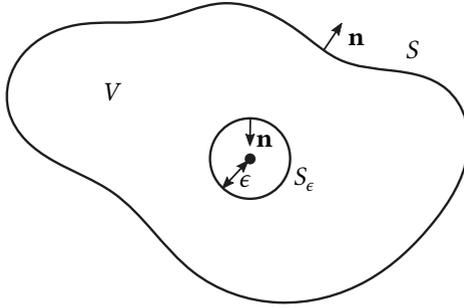


Figure A.1: Region of for the Helmholtz-Kirchhoff integral.

the volume V is composed of two parts, which are the outer surface S and the surface of the infinitesimal sphere S_ϵ . Within this region of integral, the singularity of $G(\mathbf{r}, \mathbf{r}')$ has been removed, and Equation A.3 becomes

$$\nabla^2 G(\mathbf{r}, \mathbf{r}') + k^2 G(\mathbf{r}, \mathbf{r}') = 0, \text{ for } \mathbf{r}' \in V. \quad (\text{A.5})$$

Therefore, $u(\mathbf{r})$ and $G(\mathbf{r}, \mathbf{r}')$ can be successfully substituted into Equation A.1, which leads to

$$\begin{aligned} & \iiint_V u(\mathbf{r}') \nabla^2 G(\mathbf{r}, \mathbf{r}') - G(\mathbf{r}, \mathbf{r}') \nabla^2 u(\mathbf{r}') d^3 r' \\ &= \oiint_{S+S_\epsilon} u(\mathbf{r}') \frac{\partial G(\mathbf{r}, \mathbf{r}')}{\partial \mathbf{n}} - G(\mathbf{r}, \mathbf{r}') \frac{\partial u(\mathbf{r}')}{\partial \mathbf{n}} d^2 r'. \end{aligned} \quad (\text{A.6})$$

Substituting Equation A.2 and A.5 into Equation A.6 apparently lead to the left half of the above equation to be zero. Therefore, the rest of the equation can be rewritten as

$$\begin{aligned} & \iint_S u(\mathbf{r}') \frac{\partial G(\mathbf{r}, \mathbf{r}')}{\partial \mathbf{n}} - G(\mathbf{r}, \mathbf{r}') \frac{\partial u(\mathbf{r}')}{\partial \mathbf{n}} d^2 r' \\ &= - \iint_{S_\epsilon} u(\mathbf{r}') \frac{\partial G(\mathbf{r}, \mathbf{r}')}{\partial \mathbf{n}} - G(\mathbf{r}, \mathbf{r}') \frac{\partial u(\mathbf{r}')}{\partial \mathbf{n}} d^2 r'. \end{aligned} \quad (\text{A.7})$$

According to the expression of $G(\mathbf{r}, \mathbf{r}')$ in Equation A.4, its derivative to the normal \mathbf{n} is calculated as

$$\frac{\partial G(\mathbf{r}, \mathbf{r}')}{\partial \mathbf{n}} = \cos \angle(\mathbf{n}, \mathbf{r} - \mathbf{r}') \left(ik - \frac{1}{|\mathbf{r} - \mathbf{r}'|} \right) \frac{e^{ik|\mathbf{r} - \mathbf{r}'|}}{|\mathbf{r} - \mathbf{r}'|}, \quad (\text{A.8})$$

where $\angle(\mathbf{n}, \mathbf{r} - \mathbf{r}')$ denotes the angle between the two vectors \mathbf{n} and $\mathbf{r} - \mathbf{r}'$. Now considering the surface S_ϵ is a sphere with the radius ϵ around the point \mathbf{r} , the angle $\angle(\mathbf{n}, \mathbf{r} - \mathbf{r}')$ is always π , and Equation A.4 and A.8 become

$$G(\mathbf{r}, \mathbf{r}') = \frac{e^{ik\epsilon}}{\epsilon}, \quad (\text{A.9})$$

$$\frac{\partial G(\mathbf{r}, \mathbf{r}')}{\partial \mathbf{n}} = \left(\frac{1}{\epsilon} - ik \right) \frac{e^{ik\epsilon}}{\epsilon}. \quad (\text{A.10})$$

Then, for the right side of Equation A.7, since the sphere S_ϵ is infinitesimal, $u(\mathbf{r}')$ and $\partial u(\mathbf{r}')/\partial \mathbf{n}$ can be seen as invariant and equal to $u(\mathbf{r})$ and $\partial u(\mathbf{r})/\partial \mathbf{n}$. By further substituting $\epsilon \rightarrow 0$, most components can be eliminated and the right side of Equation A.7 can be simplified as

$$\begin{aligned} \iint_{S_\epsilon} u(\mathbf{r}') \frac{\partial G(\mathbf{r}, \mathbf{r}')}{\partial \mathbf{n}} - G(\mathbf{r}, \mathbf{r}') \frac{\partial u(\mathbf{r}')}{\partial \mathbf{n}} d^2 r' \\ = 4\pi\epsilon^2 \left[u(\mathbf{r}) \left(\frac{1}{\epsilon} - ik \right) \frac{e^{ik\epsilon}}{\epsilon} - \frac{\partial u(\mathbf{r})}{\partial \mathbf{n}} \frac{e^{ik\epsilon}}{\epsilon} \right] \\ = 4\pi u(\mathbf{r}), \end{aligned} \quad (\text{A.11})$$

Therefore, the following equation can be derived from Equation A.7

$$u(\mathbf{r}) = \frac{1}{4\pi} \iint_S G(\mathbf{r}, \mathbf{r}') \frac{\partial u(\mathbf{r}')}{\partial \mathbf{n}} - u(\mathbf{r}') \frac{\partial G(\mathbf{r}, \mathbf{r}')}{\partial \mathbf{n}} d^2 r', \quad (\text{A.12})$$

or by substituting the Green's function that Kirchhoff has chosen, it can be written as

$$u(\mathbf{r}) = \frac{1}{4\pi} \iint_S \frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} \frac{\partial u(\mathbf{r}')}{\partial \mathbf{n}} - u(\mathbf{r}') \frac{\partial}{\partial \mathbf{n}} \left(\frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} \right) d^2r'. \quad (\text{A.13})$$

The above equation is known as the Helmholtz-Kirchhoff integral theorem, which plays an important role in the development of the diffraction theory. It implies that the field at any point within a volume can be calculated by knowing the field distribution and its derivative to the normal direction of the surrounding surface. However, in reality, usually only the field on a small part of the surface, e.g. an aperture, is known.

Thus, with the above theorem, Kirchhoff has worked further to solve the problem of diffraction of an aperture as shown in Figure A.2. The integration surface is divided into two parts, i.e., S_1 which is the aperture plane and S_2 which is an infinitely large sphere around the point \mathbf{r} . Besides, the region of the aperture is denoted as Σ , which is located on S_1 .

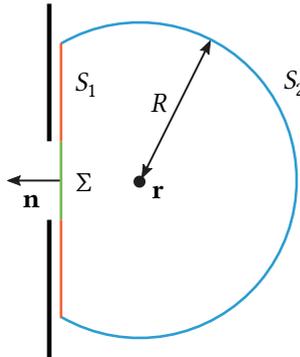


Figure A.2: Kirchhoff diffraction formulation of an aperture on a plane.

With the above model, Equation A.12 is also divided into two integrals on S_1 and S_2 respectively. Then the field at point \mathbf{r} becomes

$$u(\mathbf{r}) = \frac{1}{4\pi} \iint_{S_1+S_2} G(\mathbf{r},\mathbf{r}') \frac{\partial u(\mathbf{r}')}{\partial \mathbf{n}} - u(\mathbf{r}') \frac{\partial G(\mathbf{r},\mathbf{r}')}{\partial \mathbf{n}} d^2r', \quad (\text{A.14})$$

Starting with S_2 , since the surface area has an order of R^2 , the way to simplify the integral is to first eliminate all the components which decay faster than R^{-2} as $R \rightarrow \infty$. For the Green's function, it is easy to know that its modulus as

$$G(\mathbf{r},\mathbf{r}') = \left| \frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} \right| = \left| \frac{e^{ikR}}{R} \right| = \frac{1}{R}. \quad (\text{A.15})$$

Next, for the derivative of the Green's function to the normal direction, it can be calculated with the Taylor expansion as

$$\frac{\partial G(\mathbf{r},\mathbf{r}')}{\partial \mathbf{n}} = \left(ik - \frac{1}{R} \right) \frac{e^{ikR}}{R} = ikG(\mathbf{r},\mathbf{r}') + \mathcal{O}(R^{-2}), \quad (\text{A.16})$$

where all the terms in $\mathcal{O}(R^{-2})$ can be omitted if $u(\mathbf{r})$ is assumed to decay like a spherical wave with the order of R^{-1} . Therefore, with the above analysis, the surface integral on S_2 in Equation A.14 can be expressed by integral on a sphere in terms of the solid angle Ω as the following equation shows

$$\begin{aligned} & \iint_{S_2} G(\mathbf{r},\mathbf{r}') \frac{\partial u(\mathbf{r}')}{\partial \mathbf{n}} - u(\mathbf{r}') [ikG(\mathbf{r},\mathbf{r}') + \mathcal{O}(R^{-2})] d^2r' \\ &= \iint_{S_2} \left\{ \left[\frac{\partial u(\mathbf{r}')}{\partial \mathbf{n}} - ik u(\mathbf{r}') \right] G(\mathbf{r},\mathbf{r}') - u(\mathbf{r}') \mathcal{O}(R^{-2}) \right\} R^2 d\Omega \\ &= \iint_{S_2} \left[\frac{\partial u(\mathbf{r}')}{\partial \mathbf{n}} - ik u(\mathbf{r}') \right] R d\Omega, \quad \text{for } R \rightarrow \infty. \end{aligned} \quad (\text{A.17})$$

Obviously, the whole integral in the above equation can be eliminated by the following assumption

$$\lim_{R \rightarrow \infty} \left[\frac{\partial u(\mathbf{r}')}{\partial \mathbf{n}} - iku(\mathbf{r}') \right] R = 0, \quad (\text{A.18})$$

which is known as the Sommerfeld radiation condition. With this assumption, the total integral in Equation A.14 is reduced to the surface integral on S_1 solely. Kirchhoff further made an assumption that the field distribution $u(\mathbf{r}')$ and its derivative $\partial u(\mathbf{r}')/\partial \mathbf{n}$ are zero across the whole plane S_1 except on the aperture Σ , and they remain exactly the same with or without the screen, which are known as the Kirchhoff boundary conditions. Therefore, the field distribution at any point \mathbf{r} behind the aperture can be solely calculated by the field distribution $u(\mathbf{r}')$ on the aperture

$$u(\mathbf{r}) = \frac{1}{4\pi} \iint_{\Sigma} G(\mathbf{r}, \mathbf{r}') \frac{\partial u(\mathbf{r}')}{\partial \mathbf{n}} - u(\mathbf{r}') \frac{\partial G(\mathbf{r}, \mathbf{r}')}{\partial \mathbf{n}} d^2 r', \quad (\text{A.19})$$

which is the formula of Kirchhoff's diffraction theory.

**SCHRIFTENREIHE AUTOMATISCHE SICHTPRÜFUNG UND BILDVERARBEITUNG
(ISSN 1866-5934)**

- Band 1 **JONATHAN BALZER**
Regularisierung des Deflektometrieproblems Grundlagen
und Anwendung.
ISBN 978-3-86644-230-6
- Band 2 **IOANA GHEȚA**
Fusion multivariater Bildserien am Beispiel eines Kamera-Arrays.
ISBN 978-3-86644-684-7
- Band 3 **STEFAN BRUNO WERLING**
Deflektometrie zur automatischen Sichtprüfung
und Rekonstruktion spiegelnder Oberflächen.
ISBN 978-3-86644-687-8
- Band 4 **JAN WASSENBERG**
Efficient Algorithms for Large-Scale Image Analysis.
ISBN 978-3-86644-786-8
- Band 5 **MARTIN GRAFMÜLLER**
Verfahrensfortschritte in der robusten Echtzeiterkennung
von Schriftzeichen.
ISBN 978-3-86644-979-4
- Band 6 **JÜRGEN BRAUER**
Human Pose Estimation with Implicit Shape Models.
ISBN 978-3-7315-0184-8
- Band 7 **MARKUS MÜLLER**
Szeneninterpretation unter Verwendung multimodaler Sensorik
und Salienzmaßen.
ISBN 978-3-7315-0240-1
- Band 8 **ROBIN GRUNA**
Beleuchtungsverfahren zur problemspezifischen Bildgewinnung
für die automatische Sichtprüfung.
ISBN 978-3-7315-0313-2
- Band 9 **THOMAS STEPHAN**
Beitrag zur Unterwasserbildrestauration.
ISBN 978-3-7315-0579-2

- Band 10 **JAN-PHILIP JARVIS**
A Contribution to Active Infrared Laser Spectroscopy
for Remote Substance Detection.
ISBN 978-3-7315-0725-3
- Band 11 **MIRO TAPHANEL**
Chromatisch konfokale Triangulation – Hochgeschwindigkeits 3D-Sensorik
auf Basis der Wellenlängenschätzung mit optimierten Filtern.
ISBN 978-3-7315-0646-1
- Band 12 **SEBASTIAN HÖFER**
Untersuchung diffus spiegelnder Oberflächen mittels
Infrarotdeflektometrie.
ISBN 978-3-7315-0711-6
- Band 13 **MATTHIAS RICHTER**
Über lernende optische Inspektion am Beispiel der
Schüttgutsortierung.
ISBN 978-3-7315-0842-7
- Band 14 **MATHIAS ZIEBARTH**
Wahrnehmungsgrenzen kleiner Verformungen auf
spiegelnden Oberflächen.
ISBN 978-3-7315-0890-8
- Band 15 **JOHANNES MEYER**
Light Field Methods for the Visual Inspection of Transparent Objects.
ISBN 978-3-7315-0912-7
- Band 16 **MASOUD ROSCHANI**
Probabilistische Planungsverfahren für die
deflektometrische Oberflächeninspektion.
ISBN 978-3-7315-0907-3
- Band 17 **MAHSA MOHAMMADIKAJI**
Simulation-based Planning of Machine Vision Inspection Systems
with an Application to Laser Triangulation.
ISBN 978-3-7315-0989-9
- Band 18 **DING LUO**
High-speed surface profilometry based on an adaptive microscope
with axial chromatic encoding.
ISBN 978-3-7315-1061-1
- Band 19 **ZHENG LI**
Application of diffractive lens arrays in confocal microscopy.
ISBN 978-3-7315-1188-5

Confocal microscopy has been the gold standard for surface measurement and life science for years. In traditional confocal microscopes, high resolution and large fields of view are difficult to achieve simultaneously by the microscope objectives. To solve the problem, diffractive lens arrays (DLAs) with overlapping apertures have been proposed in previous research. They can produce spots with high numerical apertures in a dense grid over a large area. The tiny spots can be used to scan the whole sample with a much higher resolution than the objectives. However, in previous research, they have only been demonstrated in transmitted-light microscopes, which significantly limits their applications.

In this work, two new DLA concepts are proposed for applications in reflected-light microscopes. The first one is designed by the superposition of a plane wave component to increase the zero-order diffraction efficiency. The second one is designed by the superposition of a second lens component and it acts exactly as an array of finite-conjugate objectives. Based on the proposed concepts, prototypes are manufactured, and experiments are carried out to validate their performance. Both kinds of DLAs have shown capabilities of high-resolution measurement over a large area. They have overcome the limitations in the previous research and realized multi-spot confocal imaging in surface metrology and fluorescence microscopy.

ISSN 1866-5934

ISBN 978-3-7315-1188-5

Gedruckt auf FSC-zertifiziertem Papier

