

Ding Luo

High-speed surface profilometry based on an adaptive microscope with axial chromatic encoding



Schriftenreihe Automatische Sichtprüfung und Bildverarbeitung | Band 18

Ding Luo

**High-speed surface
profilometry based on an
adaptive microscope with
axial chromatic encoding**

Schriftenreihe Automatische Sichtprüfung und Bildverarbeitung
Band 18

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

Lehrstuhl für Interaktive Echtzeitsysteme
am Karlsruher Institut für Technologie

Fraunhofer-Institut für Optronik, Systemtechnik
und Bildauswertung IOSB

High-speed surface profilometry based on an adaptive microscope with axial chromatic encoding

by
Ding Luo

Karlsruher Institut für Technologie
Lehrstuhl für Interaktive Echtzeitsysteme

High-speed surface profilometry based on an adaptive microscope
with axial chromatic encoding

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

von Ding Luo

Tag der mündlichen Prüfung: 18. Dezember 2019
Erster Gutachter: Prof. Dr.-Ing. habil. Jürgen Beyerer
Zweiter Gutachter: Prof. Dr. rer. nat. Wilhelm Stork

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 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 2021 – Gedruckt auf FSC-zertifiziertem Papier

ISSN 1866-5934
ISBN 978-3-7315-1061-1
DOI 10.5445/KSP/1000125427

Abstract

For the quality assurance of a technical part, the three-dimensional (3D) geometric profile of the working surface is often one of the most important aspects, which directly affects the functionality of the part in a fundamental way. For example, the roughness of the working surface is typically under careful inspection to guarantee specific mechanical properties during its interaction with the environment or the other components. Over the past decades, optical 3D surface profilometry has gained an increasing amount of attention for such applications in both academic and industrial environments, due to its capability of non-contact measurement and high resolution. Various optical probes are designed to interact with the target surface, in order to reveal the underlying 3D structure.

With the initialization of Industry 4.0, modern “smart factories” are posing new challenges to surface profilometry technologies, demanding swift adaptation to different inspection tasks with fast measurement speed and high accuracy. Such challenges are difficult for conventional optical profilometry methods, as they are restricted by the fundamental dilemma between accuracy and speed. Technology such as the confocal scanning microscopy is celebrated for its superior resolution and accuracy while suffering from a slow measurement speed due to its requirement of the mechanical scanning as well as a low density of measurement to avoid crosstalk. On the contrary, method such as shape from focus (SFF) measures all lateral locations simultaneously, which is much more efficient. Nevertheless, the resolution and accuracy of

the measurement are degraded accordingly. In this thesis, a cascade measurement strategy is proposed for optical surface profilometry based on an adaptive microscope, which consists of a pre-measurement stage to limit the axial measurement range, a main measurement stage, and a post-measurement stage for refinement.

To realize such a strategy, an adaptive microscope with axial chromatic encoding is first designed and developed, namely the AdaScope. With a holistic design approach, the AdaScope consists of two major components. Firstly, the programmable light source is based on a supercontinuum laser, whose echellogram is spatially filtered by a digital micromirror device (DMD). By sending different patterns to the DMD, arbitrary spectra can be generated for the output light. Secondly, the programmable array microscope is constructed based on a second DMD, which serves as a programmable array of secondary light source. A chromatic objective is utilized so that the necessity of axial mechanical scanning is avoided. The combination of both components grants the AdaScope the ability to confocally address any locations within the measurement volume, which provides the hardware foundation for the cascade measurement strategy.

For the pre-measurement stage, a compressive shape from focus (CSFF) method is proposed, where the focal stack is captured in a compressive manner. Each frame is a weighted linear combination of all focal planes along the optical axis, which improves the efficiency of the capturing process. Compared to conventional SFF method, the image acquisition is 7 times faster.

Two methods are proposed for the main measurement stage. The iterative array adaptation method is based on the conventional confocal array scanning. Multiple iterations of lateral array scanning are performed for a single measurement. From iteration to iteration, the array density is increased while the axial measurement range is reduced accordingly to avoid crosstalk. Linear measurement based on two ramp illumination spectra is proposed for the axial scan to efficiently capture information regarding the surface profile.

The other candidate for the main measurement stage is direct area confocal scanning based on tilted illumination field. It is demonstrated both theoretically and experimentally that the confocal signal is largely preserved even for a wide-field illumination, as long as the illumination is tilted to a specific angle range according to the numerical aperture of the system. This leads to a much improved measurement speed with a moderately reduced sensitivity.

Last but not least, for post-measurement refinement, a dynamic sampling approach is developed based on Bayesian experimental design (BED). The calculation of the utility function involves numerical integration conducted through Monte Carlo sampling, which is computationally expensive. To accelerate the process, a recurrent neural network (RNN) is developed and trained to approximate the BED process. According to the simulation result, this approach is able to achieve a performance between uniform sampling and full BED, with a speed improvement of 600 times.

Kurzfassung

Für die Qualitätssicherung eines technischen Teils ist das dreidimensionale geometrische Profil einer Funktionsoberfläche oft einer der wichtigsten Aspekte, welcher die Funktionalität des Teils in grundlegender Weise direkt beeinflusst. Beispielsweise wird die Rauheit der Funktionsfläche normalerweise sorgfältig geprüft, um bestimmte mechanische Eigenschaften während ihrer Wechselwirkung mit der Umgebung oder anderen Bauteilen zu gewährleisten. In den letzten Jahrzehnten hat die optische 3D-Oberflächenprofilometrie aufgrund ihrer Fähigkeit zur berührungslosen Messung und hohen Auflösung für solche Anwendungen sowohl im akademischen als auch im industriellen Umfeld zunehmend an Bedeutung gewonnen. Für die Erfassung von Oberfl. Zieloberflächen wurden verschiedene optische Sonden entwickelt, um die zugrunde liegende 3D-Struktur zu messen.

Mit dem Aufkommen von Industrie 4.0 stellen moderne intelligente Fabriken neue Herausforderungen an die Oberflächenmesstechnik. Sie erfordern eine schnelle Anpassung an verschiedene Inspektionsaufgaben mit hoher Messgeschwindigkeit und hoher Genauigkeit. Solche Herausforderungen sind für herkömmliche optische Profilometrieverfahren schwierig, da sie durch das grundlegende Dilemma zwischen Genauigkeit und Geschwindigkeit begrenzt sind. Technologien wie das Konfokalmikroskop sind bekannt für ihre überlegene Auflösung und Genauigkeit, leiden aber unter einer geringen Messgeschwindigkeit, da ein mechanisches Scannen sowie eine geringe Messdichte zur Vermeidung von lateralem Übersprechen erforderlich sind. Im Gegenteil, eine Methode wie Shape from Focus misst alle benachbarten Positionen

gleichzeitig, was wesentlich effizienter ist. Allerdings verschlechtern sich Auflösung und Genauigkeit der Messung entsprechend. In dieser Arbeit wird eine Kaskadenmessstrategie für die optische Oberflächenprofilometrie vorgeschlagen, die auf einem adaptiven Mikroskop basiert und aus drei Messstufen besteht: einer Vormessstufe zur Begrenzung des axialen Messbereichs, einer Hauptmessstufe und einer Nachmessstufe zur Verfeinerung.

Um eine solche Strategie umzusetzen, wird zunächst ein adaptives Mikroskop mit axialer chromatischer Codierung entworfen und entwickelt, das sogenannte AdaScope. Mit einem ganzheitlichen Designansatz besteht das AdaScope aus zwei Hauptkomponenten. Erstens basiert die programmierbare Lichtquelle auf einem Weißlichtlaser, dessen Echellogramm durch ein Digital Mirror Device (DMD) räumlich gefiltert wird. Durch Senden verschiedener Muster an den DMD können beliebige Ausgangsspektrale erzeugt werden. Zweitens basiert das programmierbare Array-Mikroskop auf einer zweiten DMD, die als programmierbare Anordnung einer sekundären Lichtquelle dient. Ein chromatisches Objektiv wird verwendet, um die Notwendigkeit einer axialen mechanischen Abtastung zu vermeiden. Die Kombination beider Komponenten ermöglicht es dem AdaScope, beliebige Stellen innerhalb des Messvolumens konfokal anzusprechen, was die Hardware-Grundlage für die Kaskaden-Messstrategie bildet.

Für die Vormessphase wird eine Compressive Shape from Focus-Methode vorgeschlagen, bei der der Fokusstapel auf komprimierende Weise erfasst wird. Jeder Frame ist eine gewichtete lineare Kombination aller Fokusebenen entlang der optischen Achse, was die Effizienz des Erfassungsprozesses verbessert. Im Vergleich zur herkömmlichen Methode Shape from Focus ist die Bildaufnahme siebenmal schneller.

Für die Hauptmessstufe werden zwei Methoden vorgeschlagen. Das iterative Anordnungsanpassungsverfahren basiert auf herkömmlichem konfokalen Abtasten. Für eine einzelne Messung werden mehrere Iterationen des lateralen Array-Scannens durchgeführt. Von Iteration zu Iteration wird die Arraydichte erhöht, während der axiale Messbereich entsprechend verringert wird, um ein laterales Übersprechen zu vermeiden. Für den axialen Scan wird

eine lineare Messung basierend auf zwei Rampenbeleuchtungsspektren vorgeschlagen, um Informationen bezüglich des Oberflächenprofils effizient zu erfassen.

Der andere Kandidat für die Hauptmessstufe ist das direkte konfokale Scannen basierend auf einem geneigten Beleuchtungsfeld. Sowohl theoretisch als auch experimentell wird gezeigt, dass das konfokale Signal auch bei einer Hellfeldbeleuchtung weitgehend erhalten bleibt, solange die Beleuchtung entsprechend der numerischen Apertur des Systems auf einen bestimmten Winkelbereich geneigt wird. Dies führt zu einer deutlich verbesserten Messgeschwindigkeit bei moderat reduzierter Empfindlichkeit.

Zu guter Letzt wird zur Verfeinerung nach der Messung ein dynamischer Abtastansatz entwickelt, der auf dem Bayesian Experimental Design basiert. Die Berechnung der Nutzenfunktion beinhaltet eine numerische Integration, die durch Monte-Carlo-Abtastung angenähert wird, was jedoch rechenintensiv ist. Um den Prozess zu beschleunigen, wird ein Recurrent Neural Network entwickelt und trainiert, um den Bayesian Experimental Design-Prozess zu approximieren. Entsprechend dem Simulationsergebnis ist dieser Ansatz in der Lage, eine Leistung zwischen einheitlicher Abtastung und vollständigem Bayesian Experimental Design mit einer Geschwindigkeitsverbesserung um das 600-fache zu erzielen.

Acknowledgements

I would like to express my sincerest gratitude to Prof. Dr.-Ing. habil. Jürgen Beyerer for providing me the opportunity to work under his guidance at the Vision and Fusion Laboratory (IES) of Karlsruhe Institute of Technology (KIT, Germany) in cooperation with Fraunhofer Institute of Optronics, System Technologies and Image Exploitation (Fraunhofer IOSB, Germany). This doctoral thesis would not have been possible without his constant encouragement and support. Many thanks also go to Prof. Dr. rer. nat. Wilhelm Stork from the Institute for Information Processing Technologies (ITIV, KIT) for serving as the second reviewer and for his valuable comments and suggestions regarding my thesis. As a kind mentor, he has opened the door of optical sensing for me since he supervised my master's thesis.

The research presented in this thesis has been conducted mainly under a collaboration project with the Institute of Applied Optics (ITO) at the University of Stuttgart, which is kindly funded by Baden-Württemberg Stiftung gGmbH. I would like to thank Prof. Dr. Wolfgang Osten, Dr. Daniel Claus, Dr. Tobias Heist, and Tobias Boettcher from ITO for the smooth and fruitful collaboration. Additionally, many thanks go to Prof. Dr.-Ing. Fernando Puente León and Dr.-Ing. Sebastian Bauer from the Institute of Industrial Information Technology (IIIT, KIT) for the collaboration on optical unmixing which initiated my research on adaptive optical measurement.

I am greatly indebted to Dr.-Ing. Miro Taphanel, my former group leader at the department of Visual Inspection Systems (SPR, Fraunhofer IOSB), for offering me the opportunity to work as a Hiwi student during my master's studies and later to join his group as a doctoral researcher. His wisdom and

humor have made the past few years a most enjoyable experience. And he has been a great source of inspiration to me, both professionally and personally.

Furthermore, I would like to thank all the colleagues at IES (KIT), especially Dr. Alexey Pak, Dr.-Ing. Chengchao Qu, Dr.-Ing. Johannes Meyer, Dr.-Ing. Matthias Richter, Chia-Wei Chen, Mahsa Mohammadikaji, Ankush Meshram, Florian Becker, Julius Krause, Patrick Philipp, Mathias Anneken, and of course my office mate Zheng Li, for the fruitful discussions, the valuable advice, the days and nights of working together before deadlines, and the beers and fun we have shared in the past few years. Meanwhile, I would like to express my appreciation to the colleagues at SPR (Fraunhofer IOSB), in particular Prof. Dr.-Ing. Thomas Längle, Christian Negara, Dennis Heddendorp, Georg Maier, Kai Niedernberg, Alexander Enderle, Dr.-Ing. Robin Gruna, Dr.-Ing. Matthias Hartrumpf for the great atmosphere, friendship and support. Many thanks also go to the non-scientific staff Petra Riegel, Britta Ost and Gaby Gross. Additionally, I am deeply indebted to Dr. rer. nat. Gunnar Ritt from the department of Optronics (Fraunhofer IOSB) for kindly lending the supercontinuum laser to me, which served as the workhorse for the experiments. I feel really honored to have the privilege of working with these wonderful people.

I would like to express my appreciation to my parents for their unconditional love and support, both physically and mentally. Last but not least, I would like to thank my beautiful wife, Qian Xu. With her brightness, understanding and devotion, she has made me who I am today. I would like to dedicate this thesis to her with love and gratitude.

Karlsruhe, July 2019

Ding Luo

Contents

Notation	XV
Acronyms	XXI
1 Introduction	1
1.1 Motivation	1
1.2 Research Topics	5
1.3 Main Contributions	8
1.4 Thesis Outline	9
2 Related Work	11
2.1 Tunable Light Source	11
2.2 Shape from Focus	13
2.3 Confocal 3D Microscopy	14
2.3.1 Confocal Lateral Scanning	15
2.3.2 Confocal Axial Scanning	18
2.3.3 Confocal 3D Scanning	23
3 Design and Construction of AdaScope	25
3.1 Programmable Light Source	25
3.1.1 Design and Simulation	26
3.1.2 Setup and Alignment	35
3.1.3 Spectrum Generation	38
3.1.4 Calibration and Results	41
3.1.5 Discussion	49
3.2 Programmable Array Microscope	52
3.2.1 System Design	52

3.2.2	Simulation and Construction	54
3.2.3	Camera Calibration	56
3.2.4	Illumination Generation	58
3.2.5	Synchronization Mechanism	60
3.3	Summary	62
4	Cascade Measurement Strategy	63
4.1	Compressive Shape from Focus	65
4.1.1	Background	65
4.1.2	Linear Measurement Model	67
4.1.3	Compressive Algorithm	69
4.1.4	Simulation and Discussion	71
4.1.5	Summary	77
4.2	Iterative Array Adaptation for 3D Confocal Scanning	78
4.2.1	Motivation and Concept	78
4.2.2	Axial Measurement Refinement	79
4.2.3	Lateral Array Condensation	81
4.2.4	Triggering Mechanism	82
4.2.5	Summary	82
4.3	Direct Area Confocal Scanning	83
4.3.1	Theoretical Analysis	84
4.3.2	Scanning Mechanism	94
4.3.3	Summary	99
4.4	RNN-accelerated Experimental Design	99
4.4.1	Chromatic Confocal Signal	100
4.4.2	Bayesian Experimental Design	103
4.4.3	RNN-based Acceleration	107
4.4.4	Summary	112
5	Evaluation and Results	113
5.1	Experimental Setup	113
5.2	Benchmark: Confocal Array Scanning	115
5.3	Pre-measurement: Compressive Shape from Focus	119
5.4	Main Measurement I: Iterative Array Adaptation	124
5.5	Main Measurement II: Direct Area Scanning	131

5.6	Analysis and Comparison	135
6	Concluding Remarks	139
6.1	Conclusion	139
6.2	Outlook	141
A	Appendix	145
A.1	Camera Calibration	145
A.2	Bernstein Polynomials	146
	Bibliography	149
	Publications	163

Notation

This chapter introduces the notation and symbols which are used in this thesis.

General notation

Identifiers & Operators	Roman letters	a, A
Scalars	italic Roman letters	x, X
	italic Greek letters	α, Φ
Vectors	bold Roman and Greek lowercase letters	$\mathbf{x}, \boldsymbol{\theta}$
Matrices & Tensors	bold Roman and Greek uppercase letters	$\mathbf{A}, \boldsymbol{\Phi}$
Sets	blackboard bold Roman uppercase letters	\mathbb{D}
Distributions	calligraphic Roman uppercase letters	\mathcal{N}

Symbols

$(\cdot)^T$	transpose of a matrix
$(\cdot)^+$	Moore–Penrose pseudoinverse of a matrix
$\ \cdot\ _2$	Euclidean norm
δz	small axial deviation from the focal plane
$\Delta\lambda_{\text{FSR}}$	free spectral range
θ	tilting angle of the illumination field

θ_{IN}	incidence angle of light
θ_{OUT}^m	diffraction angle of light at a diffraction order of m
$\boldsymbol{\theta} = (\theta_1, \theta_2)$	parameters to be estimated from the confocal signal
λ	wavelength of light
λ_t	wavelength to be measured at a certain time
λ_{S}	shortest wavelength in the free spectral range
λ_{L}	longest wavelength in the free spectral range
μ	parameter of the Gaussian function
ξ	possible design for the next experiment in Bayesian experimental design
ξ^*	optimum design for the next experiment in Bayesian experimental design
σ	parameter of the Gaussian function
σ_{n}^2	variance of noise
ϕ_{B}	blazing angle of grating
$\Phi_{\text{a}}(\lambda, t)$	temporal spectral flux of the programmable light source
$\Phi_{\lambda, t}$	discrete temporal spectral flux of the programmable light source as a matrix
$A_{(\cdot)}$	factors for the variance and bias of the Monte Carlo estimator for the utility function
$a_{z, \lambda}$	axial chromatic focal shift
A	amplitude of a Gaussian function
\mathbf{A}	measurement matrix
$\mathbf{b}_{(\cdot)}$	bias of the corresponding layer in a neural network
$\text{COG}(\cdot)$	center of gravity
c	a factor to match scaling and/or unit
$C(u, v)$	component of 3D point spread function

$D_{\text{KL}}(\cdot)$	Kullback-Leibler divergence
d	grating period
d_p	physical width of a DMD pixel
d_1	object distance to the lens
d_2	image distance to the lens
D_a	lateral domain of the spectral DMD
D_λ	wavelength range of the system
$E(\cdot)$	expectation value
$E_a(x_a, y_a, \lambda)$	spectral flux density of the echellogram
$E_b(x_b, y_b, \lambda, t)$	temporal spectral flux density incident on the spatial DMD
e_i	discrete spectral flux of the i -th pixel on the spectral DMD
$E_a, E_{a'}$	discrete spectral flux density of the echellogram as a matrix
E_{3D}	discrete spectral flux density of the echellogram as a 3D tensor
f	focal length
$f(\cdot)$	activation function of a neural network layer
$F(\cdot)$	3D mask function for the illumination distribution
f_{LAP}	1D Laplacian filter
g	measurement at a certain wavelength
g_t	measurement at a certain wavelength at a certain time
\hat{g}	expectation of the measurement
G	a collection of all current measurements
$h(u, v)$	amplitude point spread function
$H(\cdot)$	intensity point spread function
$H_\lambda(x, y, z, \lambda)$	normalized chromatic intensity point spread function
$I(\cdot)$	intensity response to a point object
$I_{\text{int}}(\cdot)$	integrated intensity response

$J_n(\cdot)$	Bessel function of first kind of order n
k	wave number
\mathbf{k}_t	LSTM layer in a neural network
l	1D index of DMD pixels
\mathbf{l}_t	hidden neural network layer to encode the measurement location
m	diffraction order
\mathbf{m}_t	hidden neural network layer to encode the measured intensity
M	parameter of Monte Carlo sampling for the utility function
$M_{(\cdot)}$	magnification of the optical system
n_a	number of effective spectral DMD pixels
n_b	number of spatial DMD pixels
n_p	array pitch distance in terms of DMD pixels
n_t	number of time steps
n_z	number of axial positions
N	parameter of Monte Carlo sampling for the utility function
\mathcal{N}	Gaussian distribution
\mathbf{o}	output layer of a neural network
$p(\cdot)$	probability density
$Q(\lambda)$	spectral energy distribution of the programmable light source
\mathbf{q}	discrete spectral energy distribution of the programmable light source as a vector
r	radial coordinate
$r_a(x_a, y_a, t)$	temporal reflectivity of the spectral DMD
$r_b(x_b, y_b, t)$	temporal reflectivity of the spatial DMD

$R_a(x_a, y_a)$	average reflectivity of the spectral DMD
r_a	discrete average reflectivity of the spectral DMD
\mathbf{R}_a	discrete reflectivity of the spectral DMD as a matrix
\mathbf{R}_b	discrete reflectivity of the spatial DMD as a matrix
\mathbb{R}	real numbers
$\sin \alpha$	numerical aperture
s	hidden layer of a neural network
$S(u, v)$	component of 3D point spread function
S_{I}	area of the homogenized illumination from the programmable light source
t	time
T_a	integration time of the spectral DMD
T_b	integration time of the spatial DMD
$U(\cdot)$	utility for Bayesian experimental design
u	axial optical coordinate
$U(\cdot)$	illumination distribution of the AdaScope
\mathbf{U}	discrete illumination distribution of the AdaScope as a matrix
v	radial optical coordinate
$\mathbf{W}_{(\cdot)}$	weight of the corresponding layer in a neural network
\mathbf{W}_k	a collection of weight matrices for the LSTM layer in a neural network
x, y, z	world coordinates
x_a, y_a	spectral DMD coordinates
x_b, y_b	spatial DMD coordinates
X_b, Y_b	spatial DMD pixel coordinates
x_c, y_c	camera coordinates

x_i	the i -th component of the signal vector
\mathbf{x}	signal vector
\mathbf{x}_r	reconstructed signal vector
y_i	the i -th component of the measurement vector
\mathbf{y}	measurement vector

Acronyms

2D	two-dimensional
3D	three-dimensional
AOL	acousto-optic lens
AOTF	acousto-optic tunable filter
BED	Bayesian experimental design
CAD	Computer-aided Design
CCD	charge-coupled device
CCSI	chromatic confocal spectral interferometry
CLS	confocal line scan
COG	center of gravity
CSFF	compressive shape from focus
DLP	Digital Light Processing
DMD	digital micromirror device
DOE	diffractive optical element
EOL	electro-optic lens
FoV	field of view
FWHM	full width at half maximum
HDMI	High-Definition Multimedia Interface

IES	German: Lehrstuhl für Interaktive Echtzeitsysteme, Vision and Fusion Laboratory
IOSB	German: Fraunhofer-Institut für Optronik, Systemtechnik und Bildauswertung, Fraunhofer Institute of Optronics, System Technologies and Image Exploitation
KIT	Karlsruhe Institute of Technology
KL	Kullback-Leibler
LAPD	diagonal Laplacian operator
LAPM	modified Laplacian operator
LCoS	liquid crystal on silicon
MC	Monte Carlo
MCMC	Markov-Chain Monte Carlo
MEMS	micro-electro-mechanical systems
N/A	not applicable
NA	numerical aperture
ND	neutral density
PAM	programmable array microscope
PCA	principle component analysis
PDLC	polymer-dispersed liquid crystal
PSF	point spread function
RNN	recurrent neural network
sCMOS	scientific complementary metal-oxide-semiconductor
SFF	shape from focus
SFIL	steerable filters algorithm
SLM	spatial light modulator
SNR	signal-to-noise ratio

SVGA	Super Video Graphics Array, equivalent to a resolution of 800×600
TENG	Tenegrad algorithm
WDM	wavelength division multiplexing

1 Introduction

1.1 Motivation

As a high-tech strategy first launched by the German government around 2012, Industry 4.0 aims to realize a “smart factory” by combining automation and data exchange in manufacturing technologies, where computerization of manufacturing is promoted. To achieve a self-optimizing production environment, great demands have been placed upon advanced sensors for quality monitoring. Such a flexible production system requires individual inspection tasks to be solved within the production cycle. And due to the desired vast individuality of the manufactured products, statistical quality assessment by means of random sampling is no longer sufficient. A “smart measurement machine” is thus needed to swiftly adapt to different inspection tasks on-site.

Out of the various properties of a technical part, the three-dimensional geometric profile of the working surface is often one of the most important aspects for quality assurance, which directly affects the functionality of the product in a fundamental way. For example, roughness of the working surface is typically under careful inspection to guarantee specific mechanical properties during its interaction with the environment or the other components. As another example, Figure 1.1 demonstrates the measurement result of a laser welding seam using a confocal line scan (CLS) system. The surface profile of the laser welding seam directly reflects the quality of the welding process and reveals possible defects inside the welding area, which might lead to malfunction or damage, e.g., the area within the red box in the height map shows a drop of the seam height. Structural characteristics of a surface, such as step, flatness or curvature, are also common subjects for inspection.

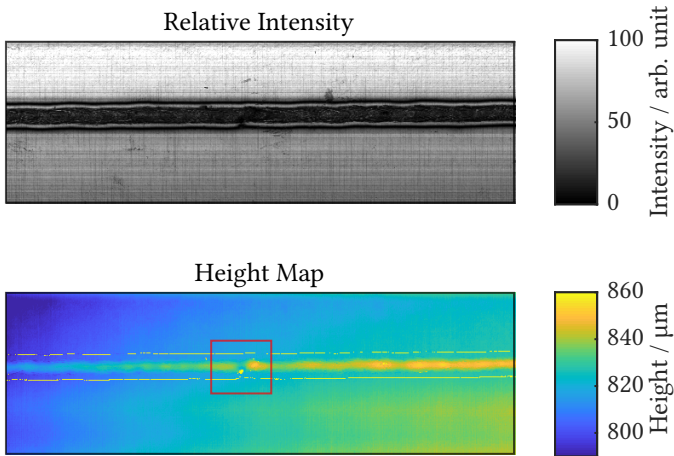


Figure 1.1: Measurement of a laser welding seam using Precitec CLS system. A defect of the welding seam has been labeled by the red box in the height map.

To solve these tasks, the conventional surface profilometer has been applied, which consists of a mechanical stylus in contact with the target under inspection. The movement of the stylus is detected and recorded as the target is scanned, which reflects its 3D profile. In recent decades, such mechanical methods have been widely replaced by optical methods, due to their advantage of non-contact measurement as well as better resolution, which are advantageous in terms of both robustness and applicability. Various kinds of optical probes are designed to interact with the target surface, in order to reveal the underlying 3D structure.

One prominent method is the confocal microscopy, which was invented by Minsky [Min61] in the late 1950s. Due to its high resolution in both the lateral direction and the axial direction, confocal microscopy has attracted much attention since the beginning. With a huge amount of research effort invested over the years, it has become a powerful tool for a wide range of applications, including scientific and industrial inspection of 3D surface profiles.

Unlike a conventional wide-field microscope, where an area illumination is applied, in a confocal system such as illustrated in Figure 1.2, a pseudo-point

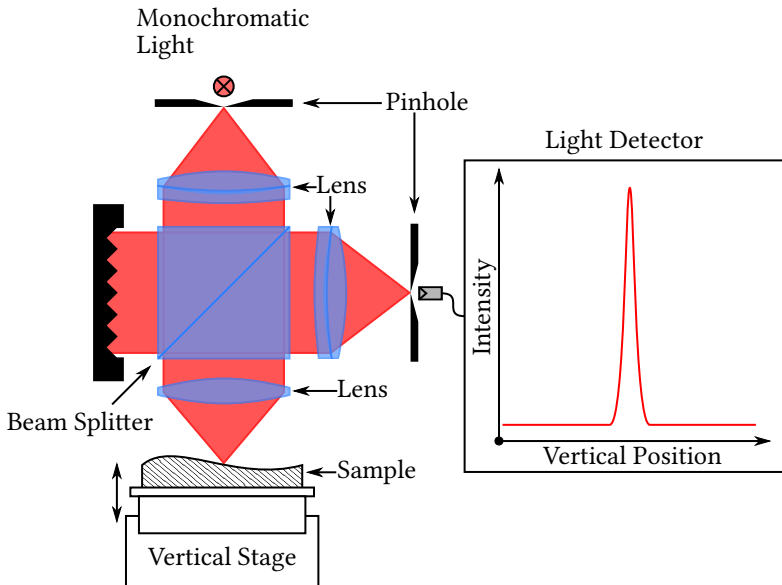


Figure 1.2: Schematic of a point confocal profile measurement system.

light source is used, which is typically achieved through filtering a normal light source with a small pinhole. Light coming out of the pinhole is focused onto the target sample surface, forming a point illumination. When the object lies exactly at the position of the focal plane, reflected light from the object surface will be able to pass through the second pinhole placed at the conjugate position in the detection arm, thus generating a high intensity value in the light detector. When the object moves away from the focal plane, the reflected light will form a blurred spot on the pinhole in front of the light detector. In this case, most of the light will be blocked by the pinhole and therefore cannot reach the light detector. By scanning the object axially in a predefined range and recording the filtered light intensity simultaneously, an intensity peak will arise, whose position indicates the height of the surface point under inspection. Through additional two-dimensional (2D) lateral scanning, the 3D profile of the target sample can be reconstructed.

With the development of digital image sensors and the advancement of computational power, various kinds of image processing techniques have been invented to retrieve information from the captured images. In a typically camera system, when the object lies within the focal plane, the corresponding image appears to be sharp with high-frequency spatial components clearly visible. However, objects out of focus will appear to be blurred, as if filtered by a low-pass filter. Based on this observation, S. K. Nayer [Nay89] first proposed the method of shape from focus in 1989. A series of images are taken, in which the focal planes of the imaging system are varied axially with respect to the target sample. For each lateral position of the object, its corresponding axial position can be retrieved by analyzing the sharpness of its adjacent area through the image stack. The image with the highest sharpness level leads to the focal plane which is the closest to the underlying lateral position. By analyzing each lateral position, the complete 3D profile of the surface can be reconstructed.

Without the necessity of lateral scanning, SFF methods are typically much faster than confocal measurement methods. However, due to the dependence on the surface texture, SFF methods are less robust compared to the confocal technologies and can only be utilized on surfaces with sufficient amount of texture. Additionally, computation of the sharpness measure requires the consideration of a sizable area adjacent to the inspected location, which effectively lowers the lateral resolution of the SFF methods.

To face the challenges presented by Industry 4.0, new measurement methods are urgently required, which can swiftly adapt to various kinds of surface profile inspection tasks. A holistic design approach must be adopted to account for different surface characteristics and scales. Fast measurement data acquisition should be coupled with advanced data processing algorithms to achieve an efficient measurement process. Additionally, the system should be able to incorporate prior knowledge regarding the product under inspection, in order to further increase the measurement speed.

1.2 Research Topics

The work presented in this thesis focuses on the advancement of microscopic surface profilometry technologies. The research problems can be categorized into the following aspects.

Optical Scanning with Minimum Mechanical Movement

The switch from a physical stylus driven by mechanical movement to an optical probe represents one of the most important advancement in the development of surface profilometry. The robustness and applicability of the measurement system is significantly improved by removing the physical contact between the measurement system and the target sample. However, macro mechanical movement between the probe and the sample is still required in most cases. For example, with a commercial chromatic line scan sensor such as the CHRocodile CLS by Precitec GmbH, at least one additional linear axis is required to achieve a complete three-dimensional measurement of a target area. For more complex situations, a two-dimensional positioning table and possibly an additional rotation stage are required to perform the measurement tasks. However, the necessity of mechanical movement presents several disadvantages.

Firstly, mechanical movement lowers the robustness of the complete measurement system. Due to the physical contacts between the components, mechanical movement systems are intrinsically more vulnerable to wear and malfunction. More maintenance effort has to be invested to ensure full functionality.

Secondly, synchronization between the measurement system and the mechanical movement further complicates system configuration. This is particularly critical for high precision microscopic measurement, since the accuracy of the measurement is directly limited by the accuracy of the mechanical movement.

Thirdly, the dependency on mechanical scanning restricts the adaptability of the measurement system on various kinds of tasks. The mechanical scanning system is typically designed and implemented for a particular measurement task. Firm connections are applied as much as possible to assure movement

accuracy. For optimum results, the complete scanning system has to be re-designed and reassembled when new tasks arise.

With the advancement of mechanical scanning devices in the past years, some of the aforementioned problems can be largely mitigated through application of state-of-the-art hardware. For example, mechanical wear can be significantly reduced through the utilization of air bearing. However, such hardware requests a significant amount of investment, which adds to the final cost of the measurement system. Meanwhile, other problems remain unsolved as long as mechanical scanning is adopted.

Therefore, this work aims to adopt a full optical scanning approach through the design and development of a novel measurement system.

Efficient Optical Information Acquisition

The development of modern computers has in many ways exceeded the improvement of electronic optical detectors in recent years. Although high speed cameras are frequently launched by manufacturers, their development speed is generally much slower compared to personal computers. Even with state-of-the-art high-speed cameras, the transfer of the image data poses new challenges on the bandwidth of communication, which limits the speed of measurement. A faster measurement system demands more efficient information acquisition and retrieval methods. In the era of digital imaging, this is equivalent to reducing the number of images needed for a certain measurement task. This requires a thorough investigation into the fundamental problems of the existing methods. For 3D surface measurement, this consists of two aspects.

In the lateral directions perpendicular to the optical axis, the most efficient way to accelerate the measurement speed is to place a magnitude of point measurement devices as densely as possible so that these locations can be measured simultaneously. However, for confocal methods, the measurement relies on the blurring of light when object is out of focus and is therefore intrinsically vulnerable to crosstalk when two measurement points are placed too close to each other. Sheppard and Mao [She88] have demonstrated the

possibility of a slit scanning confocal system through their theoretical analysis. Although infinitely dense in one lateral direction, only one location can be measured in the other direction. And for a 2D grid of confocal measurement points, the minimum pitch between the adjacent points have to be satisfied to avoid any crosstalk which might degrade the measurement accuracy, thus limiting the maximum lateral density of the measurement device. Even worse, a longer axial measurement range leads to a higher possibility of crosstalk. Additionally, when high numerical aperture (NA) optics are applied for increased resolution, the system also becomes more vulnerable to lateral crosstalk. This problem is also more severe when a 2D imaging sensor is applied, since the image captured and transferred contains largely unoccupied areas with little information regarding the actual measurement locations, which is highly inefficient.

As the density of a 2D confocal measurement grid increases, the gradually increased crosstalk reduces the confocal microscope to a conventional wide-field microscope, which loses its depth discerning capability. Luckily, with sufficient amount of surface texture, methods such as SFF can be applied to retrieve the 3D profile of the surface, albeit with reduced resolution and accuracy compared to the equivalent confocal setup. Nevertheless, SFF methods still suffer from an efficiency problem considering the axial scanning process, which also happens for the confocal methods. For both cases, an axial signal has to be retrieved with a Gaussian-like peak, whose position indicates the relative distance between the measurement system and the target object. Therefore, for both approaches, a stack of measurements have to be acquired while the optical probe (point or plane of focus) is scanned axially. To accurately locate the peak position of the axial signal, a uniform (equidistant) sampling approach is typically adopted. Although widely applied, such a sampling approach is highly inefficient as most of the sampled values are close to zero, which contains very little information regarding the position of the signal peak. According to estimation theory [van07], the measurement uncertainty is directly related with the gradient of the estimator, i.e., the slope of the signal in this case. Therefore, an objective with higher NA is preferred to generate a peak as narrow as possible. For a fixed measurement range, a narrower peak

requires denser sampling to locate accurately, further lowering the efficiency of the measurement process.

To alleviate the aforementioned issues, several measurement methods have been developed to enhance the efficiency of optical information acquisition in surface profilometry.

1.3 Main Contributions

This thesis aims to develop an optical system for high-speed surface profilometry with a holistic approach. The main contributions are listed below.

Firstly, an adaptive microscope with axial chromatic encoding has been designed and constructed, namely the AdaScope.

- programmable light source has been constructed based on the echellogram of a super-continuum laser and a DMD [Luo17c]. The system is capable of generating spectral peaks with a minimum full width at half maximum (FWHM) of less than 1 nm. When acting as a scanning bandpass filter, the wavelength tuning resolution can reach as small as 0.01 nm.
- A programmable array microscope has been proposed, which is coupled with the programmable light source to allow for the generation of arbitrary 3D illumination field [Luo18].

Based on the AdaScope platform, various 3D measurement principles have been proposed.

- A compressive shape from focus method has been proposed where the focal stack is compressively captured and the focus measure is reconstructed computationally [Luo16a, Luo16c].
- A confocal array scanning principle has been proposed where the axial measurement range and the lateral array density are adapted iteratively [Luo18].

- A confocal area scanning principle has been proposed based on a tilted illumination field [Luo20]. Theoretical analysis has been performed together with numerical simulation.
- For localized confocal scan, a method based on Bayesian experimental design has been proposed to improve the scanning efficiency. To accelerate the computation, a recurrent neural network has been developed and trained to approximate the process [Luo17a].

To validate the proposed system and methods, experimental evaluations have been conducted.

- Performance of the programmable light source has been experimentally validated [Luo17c] and applied in optical unmixing [Luo16b].
- The Compressive Shape from Focus method has been implemented both in simulation and in practice, which is capable of generating raw 3D measurement with a minimum number of frames [Luo16c, Luo17b].
- The iterative array adaptation method has been implemented in practice and evaluated through a series of experiments [Luo18].
- The direct area scanning method has been implemented in practice and evaluated through a series of experiments [Luo20].
- The RNN-accelerated Bayesian experimental design method has been evaluated through simulation [Luo17a].

1.4 Thesis Outline

This thesis is structured as follows:

Chapter 2: Related Work This chapter offers an extensive survey of the existing literature within the scope of this thesis, including tunable light source, the shape from focus technology and various confocal scanning methods.

Chapter 3: Design and Construction of AdaScope This chapter demonstrates the design of an adaptive microscope with axial chromatic encoding,

namely the AdaScope. The AdaScope system mainly consists of two components, the programmable light source and the programmable array chromatic microscope. Construction of each component is described in detail in their respective section.

Chapter 4: Cascade Measurement Strategy Based on the AdaScope platform presented previously, Chapter 4 discusses a cascade measurement strategy, which is formed by a series of measurement methods. The compressive shape from focus method is capable of rough measurement with a minimum number of frames, which is suitable as a pre-measurement step to limit the axial measurement range. Based on the result from the pre-measurement step, the main measurement can be initialized, where two candidate methods are presented and analyzed, i.e., iterative array adaptation and direct area scanning with tilted illumination. Last but not least, a post-measurement refinement step has been proposed based on Bayesian experimental design, which can be further accelerated through a recurrent neural network.

Chapter 5: Evaluation and Results In this chapter, a series of experiments have been implemented to evaluate the proposed system as well as the measurement methods.

Chapter 6: Concluding Remarks Lastly, the result of this thesis is summarized while an outlook of future research is also presented.

2 Related Work

The goal of this work is to construct an adaptive microscope with chromatic encoding, based on which a cascade of optical measurement methods can be developed and applied for surface profilometry. To provide a background, an extensive overview of the existing literature is presented in this chapter. These previous works are divided into three sections, which are related to different aspect of the thesis. Section 2.1 focuses on the development of a particular type of light source, whose output spectrum is tunable. As a vital component of the proposed system, the tunable light source is responsible for the axial scanning of the optical information acquisition process. Section 2.2 gives a brief overview of the shape from focus technology. Based on a wide field microscope, this technology achieves 3D measurement of the target surface using a stack of images, which are acquired while the focal plane is shifted with respect to the target sample. Although relatively fast in terms of image acquisition, robustness of the shape from focus technologies depends strongly on the image processing techniques as well as the intrinsic textures of the object surface. Section 2.3 explores the various confocal scanning methods in both the lateral direction and the axial direction. Compared to shape from focus technologies, confocal measurement methods are typically capable of higher resolution and accuracy while demanding a longer scanning process.

2.1 Tunable Light Source

The ability to modulate the spectrum of light serves as a powerful tool in various fields of research, including biomedical optics, optical communication, hyperspectral imaging, optical measurement, etc. The variability of the

spectrum in such systems enables various analog signal processing methods which greatly improves the performance and generates new possibilities. For example, Hirai et al. developed a multispectral image projector using a programmable spectral light source [Hir16]. By using multiple primary colors, the system is capable of wide-gamut projection.

Earlier development of spectrum modulation is mainly driven by the need for wavelength division multiplexing (WDM) in optical communication as well as chemical analysis. Various technologies have been proposed and implemented to realize a tunable spectral filter, including acousto-optics [Hua96, Bei04, Sap02], liquid crystals [Par98, Pat91], fiber bragg grating [Inu01], interferometer [Din11], etc. These methods typically focus on the realization of tunable bandpass filters, some of which with tunable bandwidth. Nevertheless, they lack the capability of manipulating a complete spectrum.

With the wide popularity of commercial Digital Light Processing (DLP) projectors, the digital micromirror device has been receiving an increasing amount of attention for the development of novel optical systems. Riza et al. introduced a digitally controlled multiwavelength programmable attenuator using a two-dimensional DMD [Riz99]. Based on this concept, a broadband optical equalizer was developed later [Riz03]. Chuang et al. proposed a programmable light spectrum synthesis system which used collimated output from a single mode fiber as the primary source [Chu06]. The light is dispersed onto a DMD chip with a diffraction grating. One dimension of the DMD is used for wavelength selection while the other dimension is used for intensity modulation of the corresponding wavelength. This concept has been commercialized using a Xenon arc lamp as the primary source and DMD for spectral filtering (OneLight Spectra by OneLight Corp. and OL-490 Agile Light Source by Optronic Laboratories). Although having been applied in surgical and biomedical research, these sources typically suffer from performance limitations due to the compromise between spectral resolution and efficiency. Wood et al. proposed to use a supercontinuum laser (also known as white light laser) as the primary source coupled with a prism as the disperser and a DMD as the modulator to construct a tunable light source [Woo12]. Diffraction analysis is made considering the DMD as a

blazed grating. The system is capable of producing illumination bands with a roughly constant width of 6 nm.

2.2 Shape from Focus

Depth estimation based on an imaging system has been a widely studied topic in the area of computer vision and image processing. Generally, existing methods can be classified into active methods and passive methods. Active methods involve the projection of an optical probe onto the target scene, often in the form of a laser beam or an illumination pattern [Sch11]. The 3D profile of the target scene is reconstructed with the information in the scattering/reflection of the optical probe captured by the imaging system. The requirement of the additional projection/illumination system will increase the complexity and the cost of the active methods, inevitably limiting their applicability. In situations where physical interaction with the scene is not allowed, passive methods are applied by taking images of the scene without specific additional illumination. Various depth cues in the captured images have been proposed by researchers, including stereopsis [Mar76], shading [Zha99], focus [Nay94], etc., which are used to reconstruct the 3D information. In this section, the usage of focus as a cue for depth measurement is discussed.

Research focuses in this area are mainly placed upon two topics, i.e., the design of robust focus measure operators and the development of estimation algorithms. Pertuz et al. [Per13] made an extensive survey and comparison of popular focus measure operators for shape from focus. Apart from the operators listed in the above survey, more complex operators are being developed constantly not only for shape from focus methods but also for sharpness estimation as a more general topic, such as the S_3 operator by Vu et al. [Vu12], which utilizes both spatial and spectral information in color images.

Conventional estimation algorithms involve localizing the maximum focus position from the focal stack for each pixel. A widely accepted method is to take a Gaussian model as proposed by Nayar et al. [Nay94]. Alternatively, other fitting methods have also been studied, such as quadratic

and polynomial fits [Sub95]. With the recent development of machine learning and optimization algorithms, more sophisticated methods have been proposed by breaking the isoplanatic restriction [Sub95], such as surface fitting and optimization by neural networks [Asi01], and total variation regularization [Mah13]. Through adoption of a deep neural network, Hazirbas et al. [Haz18] proposed a SFF method, which provides an end-to-end solution for depth reconstruction from the focal stack. As can be seen from the listed literature, the design of the focus measure operator and the development of estimation algorithm are often conducted simultaneously in a holistic manner in order to improve the performance of the overall method.

Unlike shape from defocus techniques, where the blur kernel is assumed known (e.g., [Fav05]), SFF techniques generally require a minimum number of image samples along the focal axis in order to perform robust estimation. This is realized by either shifting the focal plane or changing the relative distance between the camera and the scene, while images are captured. When a large number of images is required, such shift/movement commonly leads to a slow measurement speed and bulky systems. Additionally, the large number of the images, which is needed for evaluation, adds to the data transfer and the computational cost. Despite the development of various image processing techniques for depth reconstruction in the previous decades, the way of image capture remains relatively unchanged.

2.3 Confocal 3D Microscopy

First invented by Minsky in the late 1950s [Min61], confocal microscopy differs from conventional wide-field microscopy by the fact that both the light source and the detector are filtered by a pinhole. Such confocal filtering not only improves the lateral resolution of the microscopic system, but also enables the system to be sensitive to the axial position of the sample. As a non-contact measurement technology, confocal microscopy has been widely applied in various fields including surface profilometry, biomedical imaging as well as other applications, due to its unique capability of depth discerning. Nevertheless, to achieve complete 3D measurement of the target, confocal

microscopy in its raw form depends on the scanning of a single focused point both laterally and axially, which leads to a relatively slow measurement speed. To tackle this problem, a great amount of research effort has been invested in the past decades to accelerate the measurement speed of confocal microscopy.

2.3.1 Confocal Lateral Scanning

In the lateral direction, multiple methods have been invented to accelerate the scanning speed, which can be categorized into two different approaches: faster scanning of a single focal point and scanning of an array of focal points.

Single Point Scanning

Acceleration of single point scanning is realized through utilization of faster mechanical components. Conventional confocal scanning systems aim to generate a relative movement between the focal spot and the target sample either through physical movement of the system/sample, or through manipulation of the focal spot. With the development of opto-mechanical components, the speed of conventional scanning methods have also been improved over the years. Arrasmith et al. have developed a 2D confocal scanning system based on a MEMS bi-axial micro-mirror [Arr10]. By changing the angle of the illumination beam through controlled tilting of the micro-mirror, the focal spot of the optical system is effectively shifted. Through adoption of a micro-mirror which is electromagnetically actuated, the system is able to achieve SVGA resolution (800×600 pixels) at 56 frames per second. Similarly, Liu et al. have developed a compact fiber-optic endoscopic probe, where a MEMS scanner is integrated to achieve confocal scanning [Liu14].

Apart from scanning achieved through a single tilting mirror, a high speed laser scanning confocal microscope has been presented by Choi et al. based on the combination of a fast rotating polygonal scanning mirror for the fast axis and a bi-directional scanning galvanometer-driven mirror for the slow axis [Cho13]. The proposed system is capable of an acquisition rate up to 200 frames per second (512×512 pixels).

Based on measurement of a single focal point, methods categorized in the first approach typically hold the advantage that the light detection system and the data processing algorithms are relatively simple and straightforward. Additionally, application of MEMS components allows for miniaturization of the complete system [Liu14]. Nevertheless, measurement speed as well as robustness of such systems is largely limited by the respective mechanical components.

Array Scanning

Although the scanning of a single focal point can indeed be accelerated through application of state-of-the-art opto-mechanical components, the majority of the research effort has been focused on the second approach, which aims to achieve a multitude of measurement points simultaneously through the same optical system. Instead of scanning a single focal point, an array of focal points are illuminated, detected and shifted. Xiao et al. serve as a pioneer by proposing a confocal scanning microscope based on a spinning Nipkow disk for real-time direct viewing [Xia88]. The pinhole array on the disk performs confocal filtering for both illumination and imaging, which greatly simplifies the setup compared to earlier tandem scanning microscopes. Later, Tiziani and Uhde use a similar setup with Nipkow disk for 3D measurements [Tiz94a]. In some applications, dense lateral sampling is sometimes not necessary. Under these circumstances, a confocal sensor with a static array of measurement points is sufficient, which is denoted as a confocal matrix sensor, such as the one presented by Hillenbrand et al. [Hil15]. Apart from using a physical pinhole array, other methods have been developed to generate an array of measurement points, such as using microlens array [Tiz94b], fiber bundle [Gmi93] and diffractive optical element [Hul12]. With the development of spatial light modulators (SLMs), the idea of a mechanically shifted point array evolved into an important field of research, i.e., programmable array microscope (PAM), where various kinds of SLMs have been utilized to generate a focal point array for confocal measurement, including digital micromirror devices [Han01, Cha15], liquid crystals on silicon (LCoS) [Hag07, Kin14], and polymer-dispersed liquid crystals (PDLCs) [Cha17]. Such systems allow for dynamic control of the point

array, which enhances the speed and the flexibility of the measurement. By simultaneously measuring multiple points laterally, the measurement time of the confocal microscope is greatly reduced. However, as the numerical aperture and the target depth range increase, measurement points become more vulnerable to the crosstalk from their adjacent points, thus demanding a larger minimum pitch distance.

Line/slit confocal scanning can be seen as a special type of the array scanning method, where a continuous line of points are illuminated and imaged through a confocal system. Sheppard and Mao first presented a theoretical foundation for the slit scanning method by showing that the axial resolution degradation of slit scanning compared to a single point is well within acceptable level, especially considering the great benefit of speed acceleration [She88]. Later, Sabharwal et al. have developed a miniaturized slit-scanning confocal setup for endoscopic measurement [Sab99]. In [Poh08], Poher et al. demonstrate a slit scanning confocal microscope based on a 2D imaging sensor instead of a line detector, with which the blurred part of the slit image can also be captured. Through application of an advanced image processing technique of background subtraction, the system obtained an axial resolution even better than a point confocal system. In general, slit scanning confocal systems are more efficient than point scanning systems in terms of the scanning speed. Nevertheless, both the lateral resolution along the slit and the axial resolution are partly sacrificed. Additionally, while one lateral axis is covered by the slit, the orthogonal lateral axis still requires scanning, which is often performed mechanically.

Spectrally encoded confocal methods in the lateral direction originate from early development of 2D image transmission by WDM [Bar79, Men97]. In [Tea98], Tearney et al. have proposed a special type of slit-scanning confocal microscope based on spectral encoding where a grating is utilized to disperse light of different wavelengths onto a line. Points on the line are detected simultaneously through measurement of the returned spectrum based on Fourier-transform spectroscopy. Compared to conventional slit scanning, the spectral dispersion provides additional confocal filtering along the slit direction. Pitris et al. have conceived a novel optical design based on

two prisms and a transmission grating, enabling better miniaturization of a spectrally encoded line scanning confocal probe [Pit03]. In [Bou05], Boudoux et al. present a spectrally encoded confocal setup for 2D measurement, where one lateral axis is scanned through a rapid wavelength-swept laser and the other lateral axis is scanned with a galvanometer mounted mirror. Kim et al. have demonstrated a spectrally encoded slit confocal setup for direct 2D measurement [Kim06]. A physical slit is used to define one lateral axis, which is dispersed in the orthogonal direction so that the other lateral axis is encoded spectrally. An optical analyzer based on a 2D CCD camera is used to measure all lateral points simultaneously. In 2015, a confocal system based on the similar principle is constructed using a wavelength-swept laser and a line scan camera [Kim15]. To the author's knowledge, spectrally encoded slit confocal microscopy is currently the only method which is capable of direct dense 2D confocal measurement without the necessity of lateral scanning. Such a system typically requires very complex illumination and detection optics and axial scanning is still required for complete 3D measurement.

2.3.2 Confocal Axial Scanning

In the axial direction, improvements of confocal scanning speed are mainly achieved through two approaches, i.e., adaptive lenses and chromatic confocal technology.

Adaptive Lenses

Apart from faster linear axes used for axial scanning, various confocal systems recently utilize adaptive lenses to vary the focus of the system electronically with a very fast speed, including adaptive lens based on micro-electro-mechanical systems (MEMS), electro-optic lens (EOL) and acousto-optic lens (AOL).

Liu et al. have proposed a 3D confocal scanning microendoscope based on MEMS for both lateral and axial scanning [Liu14]. For axial scanning, a tunable MEMS lens has been constructed, which consists of a MEMS lens-scanner

with a central opening and a 2.4 mm diameter glass objective lens assembled onto the scanner platform. The central platform in the MEMS lens-scanner is symmetrically supported by four lateral-shift-free large-vertical-displacement (LSF-LVD) actuators which are driven by electrothermal actuation. In [Kou14], Koukourakis et al. have developed an adaptive lens based on Polydimethylsiloxan membrane with a piezo actuator for axial confocal scanning. With a single adaptive lens in the system, performance of the axial scan is often degraded due to defocus and spherical aberration introduced by the tuning of the adaptive lens. By employing a second adaptive lens in the detection path, aberrations are successfully balanced and homogeneous axial resolution can thus be achieved.

Based on a different principle, Shibaguchi et al. have developed a Lead-Lanthanum Zirconate-Titanate (PLZT) electro-optic lens with variable focal length [Shi92]. In [Kha06], Khan and Riza have constructed a confocal system based on a tunable focus liquid crystal lens. Despite the relative fast tuning speed, an EOL typically suffers from the drawback of being polarization sensitive.

Kaplan et al. are the first to demonstrate high-speed focus scanning using an acousto-optic lens [Kap01]. The lens consists of two adjacent acousto-optic scanners with counterpropagating acoustic waves that have the same frequency modulation but a π phase difference. Based on such a lens, a confocal profilometer is constructed, which achieves an axial scan of 400 kHz. Nevertheless, acousto-optic crystals have a weak transmission efficiency, owing to their need of multiple crystals for focusing a beam. Recently, tunable acoustic gradient (TAG) index of refraction lenses have emerged as a new generation of high-speed tunable optical components, which are able to provide complex beam profiles. Mermillod-Blondin et al. have demonstrated the use of a TAG lens as a fast varifocal element [Mer08]. Duocastella et al. have integrated an acousto-optic lens in a commercial confocal system to achieve axial scanning at 140 kHz [Duo14]. By using synchronized and high time-resolution detection, simultaneous multiplane confocal imaging can be achieved. In [Szu18], Szulzycki et al. have developed an AOL operating at a focus tuning rate of

300 kHz, which is combined with a laser scanning confocal microscope for fast 3D imaging.

Encoding of Axial Response

Axial chromatic encoding has been widely applied in confocal systems, which replaces mechanical scanning by focusing light of different wavelengths onto different axial positions. Instead of the monochromatic light source used in conventional confocal systems, chromatic confocal technology requires the usage of a polychromatic light source for illumination. Decoding of the axial information has to be conducted by a wavelength-sensitive detector such as a spectrometer or through wavelength scanning. The first chromatic confocal system is an optical profilometer presented by Molesini et al. [Mol84]. Lateral dispersion is used in the detection arm so that the spectrum of the reflected light can be measured with a photodiode array. Since then, various methods have been invented to improve the performance of chromatic confocal scanning. The application of better light sources is accompanied with new designs of the dispersion system to achieve better illumination. Advanced sampling mechanisms are developed to measure the spectral data with faster speed.

Hutley et al. have developed a wavelength-encoded linear displacement transducer based on a zone plate as the dispersing element [Hut88]. In [Tiz96], Tiziani et al. have manufactured an array of microlenses formed by zone plates to realize a chromatic confocal matrix sensor. Instead of a broadband light source, four semiconductor laser diodes with different wavelengths are used together with a CCD chip for intensity measurement. To improve upon the idea of using zone plates for dispersion, Dobson et al. have designed more complex diffractive optical elements for chromatic confocal imaging [Dob97]. Meanwhile, a tunable Ti-sapphire laser has been utilized as the light source, whose output wavelength is shifted electronically while the confocal intensity signal is recorded by the detector. Shi et al. have applied a supercontinuum light source based on non-linear effect of a photonic crystal fiber in a chromatic confocal microscope [Shi04]. Axial measurement range benefits significantly from the broad bandwidth of the supercontinuum light source.

In [Hil12b, Hil12a], Hillenbrand et al. provide comprehensive design strategies for hybrid hyperchromatic lenses based on the combination of diffractive and refractive elements. The longitudinal chromatic aberration of the system is maximized for optimum axial measurement range in a chromatic confocal system. And in [Hil13a, Hil13b], Hillenbrand et al. present a spectrally multiplexed three-point sensor using a segmented diffractive optical element (DOE). As each segment of the DOE generate different axial chromatic dispersion, a single spectrometer can be used to retrieve the axial positions of three lateral points. In [Ray14], Rayer and Mansfield have compared the performance of refractive, diffractive and hybrid aspheric diffractive optics for the application of chromatic confocal technology. The presented hybrid aspheric diffractive lens is able to combine the low geometric aberration of a diffractive lens with the high optical power of an aspheric lens, thus achieving better performance.

In the aforementioned methods, decoding of the axial information is conducted through measurement of the reflected spectrum. The spectrum is sampled either with the illumination arm or the detection arm of the system. With illumination sampling, a tunable laser source is typically used whose wavelength is scanned while a broadband detector records the reflected intensity for each wavelength. Such process is time-consuming as each wavelength has to be measured consecutively. For sampling made directly with detection, a spectrometer of a certain form is utilized to measure the reflected spectrum of a broadband illumination source. Although multiple wavelengths are measured simultaneously, readout and transfer of the spectrometer data still limit the measurement speed of the system. In both cases, dense sampling of the complete spectrum is highly inefficient, as the signal can be considered to be very sparse and the underlying parameter to be estimated is only two-dimensional, i.e., surface reflectance and axial position. The low intensity wavelength positions suffer from a low signal-to-noise ratio (SNR) mainly due to the photon noise, providing very little additional information. In practice, such noise could even have an adverse effect on a naive Gaussian fitting process. As an example, for a fiber-based chromatic confocal sensor, Luo et al. have shown that Gaussian fitting performed only for wavelengths with higher

intensity achieves a slightly higher sensitivity compared to Gaussian fitting conducted for all wavelengths [Luo12].

Therefore, to achieve more efficient sampling of the axial confocal signal, multiple methods have been developed to linearly map the high dimensional spectrum data into a lower dimensional space using spectral filters in either illumination or detection. In [Jon93], Jones and Russel have first defined the idea of wavelength discrimination systems and chromatic discrimination systems, which can be applied to chromatic confocal sensing as well. Without measuring the reflected spectrum exactly, Tiziani and Uhde have proposed a chromatic confocal microscopic system where three color filters are used with a CCD camera to generate color images of the sample [Tiz94a]. This detection setup is intrinsically equivalent to a spectrometer with very low resolution and the depth information is therefore encoded in the RGB color. This idea naturally extends to applications of multispectral and hyperspectral cameras. Kim et al. have proposed a chromatic confocal system with a wavelength detection method using transmittance [Kim13]. Two detection channels are implemented with two photomultiplier tubes, where one channel records the total reflected intensity and the other channel records the intensity of the reflected light after filtering through a color filter. In [Tap13], Taphanel et al. provide a more complex design methodology for the development of a multi-channel chromatic confocal detection system using six interference filters. Physical structure of the interference filters are directly optimized based on a merit function related to the uniqueness and sensitivity of the chromatic axial measurement.

Apart from chromatic encoding, Lee et al. have developed a confocal system based on direct mapping of axial information onto two intensity channels [Lee14], which shares some similarity to the idea of chromatic confocal detection with transmittance by Kim et al. [Kim13]. Instead of using spectral filters, two pinholes with different sizes are placed in front of the two detectors, generating two different axial response curves, which can be used to encode the axial information.

2.3.3 Confocal 3D Scanning

To perform complete 3D confocal scan, methods presented in Section 2.3.1: Confocal Lateral Scanning and Section 2.3.2: Confocal Axial Scanning are combined in a single system.

For example, Cha et al. have used chromatic axial encoding together with dynamically configurable micromirror scanning to achieve nontranslational 3D profilometry [Cha00]. In [Lin98], Lin et al. have applied chromatic encoding to a slit-scan confocal system based on a diffractive lens as the dispersing element. A 2D CCD imager is coupled with a spectral grating to achieve single-shot 3D measurement of a line. Chen et al. have used a DMD to generate a scanning array of focal points while adopting chromatic encoding for the axial measurement [Che11]. Hillenbrand et al. have presented a chromatic confocal matrix sensor based on a pinhole array [Hil13a]. And later an actuated pinhole array is adopted to achieve complete 3D scanning [Hil15]. Joeng et al. have combined a direct-view confocal microscope with an electrically tunable lens to achieve a matrix sensor for 3D scanning [Jeo16]. And in [LEE18], Lee et al. have developed a 3D confocal microscope based on a dual-detection method for axial measurement and a DMD for lateral array scanning.

3 Design and Construction of AdaScope

In this chapter, an adaptive chromatic confocal microscope is presented as the result of a holistic design approach with the target of fast 3D surface profilometry, namely the AdaScope. The AdaScope system is composed of two major components, i.e., a programmable light source and a programmable array microscope. Section 3.1 introduces the design and development of the programmable light source. Calibration and performance of the programmable spectrum generation are also described in detail. In Section 3.2, a programmable array microscope with chromatic encoding is presented. Finally, the development of the AdaScope is summarized in Section 3.3.

3.1 Programmable Light Source

Illumination systems with tunable spectrum have been receiving increasing amount of attention due to their wide applications and unique capability. In the AdaScope system, the programmable light source with fast response and accurate spectrum reproduction serves as the foundation for various axial scanning methods.

With the constant development of DMD, the binary pattern rate has been increased over the years, allowing for time multiplexing schemes to be developed for real-time applications. Therefore, the presented system utilizes the complete two-dimensional area of a DMD for wavelength selection in order to achieve superior spectral resolution, while intensity modulation is realized in a time multiplexing fashion. The system is designed and constructed based on

the orthogonally placed combination of a prism and an echelle grating to generate the echellogram of a supercontinuum laser onto the DMD. A complete calibration procedure is developed and implemented. Several spectra are generated and analyzed, indicating a minimum FWHM of less than 1 nm. When acting as a scanning bandpass filter, the wavelength tuning resolution can reach as small as 0.01 nm. The proposed filtering system can be constructed with relatively low cost and easily attached to commonly available supercontinuum lasers to generate illumination with programmable spectrum.

3.1.1 Design and Simulation

The design of the programmable light source is based on the theoretical analysis of the 2D dispersion from the prism and the echelle grating, as well as an optical simulation using the OpticStudio software by Zemax.

3.1.1.1 Theoretical Background

A diffraction grating can be characterized with the grating equation [Sos11]:

$$d(\sin \theta_{\text{OUT}}^m + \sin \theta_{\text{IN}}) = m\lambda \quad (3.1)$$

where m stands for the diffraction order, λ represents the wavelength of the light, d represents the grating period, θ_{IN} and θ_{OUT}^m represent the incidence angle and diffraction angle respectively. For a blazed grating, the blazing angle ϕ_{B} defines the angle between the facet normal and the surface normal of the grating, as shown in Figure 3.1. The blazing angle of the grating is used to redirect energy to a certain order of diffraction for better efficiency.

An echelle grating is a special type of blazed grating characterized by a large blazing angle of the grooves and used at very large diffraction orders obtaining strong dispersion. Since an echelle grating is typically installed in a Littrow configuration under blazing condition (Figure 3.2), the angular dispersion can be written as:

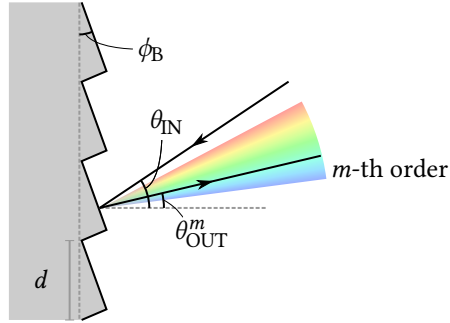


Figure 3.1: Schematic of a blazed reflection grating.

$$\frac{\partial \theta_{\text{OUT}}^m}{\partial \lambda} = \frac{m}{2d \cos \theta_{\text{OUT}}^m} = \frac{\tan \phi_B}{\lambda} \quad (3.2)$$

Although the groove density of the echelle grating is smaller than common blazed gratings, the groove structure is optimized for much larger blazing angle and therefore the light is concentrated into much higher diffraction orders. As shown in Equation 3.2, in a Littrow configuration under blazing condition, the dispersion of the grating depends only on the blazing angle of the grating and the wavelength, which allows the echelle grating to have much higher dispersion than normal blazed grating.

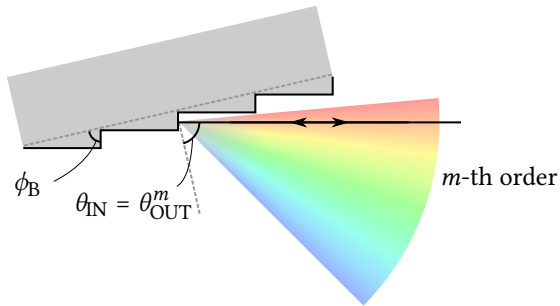


Figure 3.2: Schematic of an echelle grating under Littrow configuration.

The free spectral range $\Delta\lambda_{\text{FSR}}$ defines the largest bandwidth in one order that does not overlap with the same bandwidth in adjacent orders:

$$\Delta\lambda_{\text{FSR}} = \frac{\lambda_s}{m} = \frac{\lambda_l}{m+1} \quad (3.3)$$

where λ_s and λ_l represent the shortest and longest wavelengths in the m -th order respectively.

For an echelle grating, due to the large blazing angle, higher numbers of orders are utilized, which leads to very small free spectral range in each order. This also means that multiple orders will overlap at the same diffraction angle, making it necessary to apply a secondary disperser in the orthogonal direction in order to separate the orders from each other. Such secondary disperser, also known as a cross disperser, can be placed before or after the echelle grating.

3.1.1.2 Optical Design

The system is firstly treated as an echellogram system and designed in the sequential mode of OpticStudio. As illustrated in Figure 3.3, the echellogram system is composed of five components. A supercontinuum laser serves as the light source of the system. The target wavelength ranges from 470 nm to 700 nm. The laser beam first enters an equilateral dispersion prism made of F2 glass to be dispersed in the horizontal direction. The incidence angle with respect to the normal of the prism surface is 39° . The output beam is immediately incident on an echelle grating in the Littrow configuration under the blazing angle for vertical dispersion. The echelle grating has a blazing angle of 63.5° and a grating period of 31.6 grooves/mm. Such a combination of the prism and the echelle grating is chosen so that the echellogram covers the entire area of the DMD. Finally, the two-dimensionally dispersed light is focused onto the DMD through an achromatic doublet with a focal length of 100 mm.

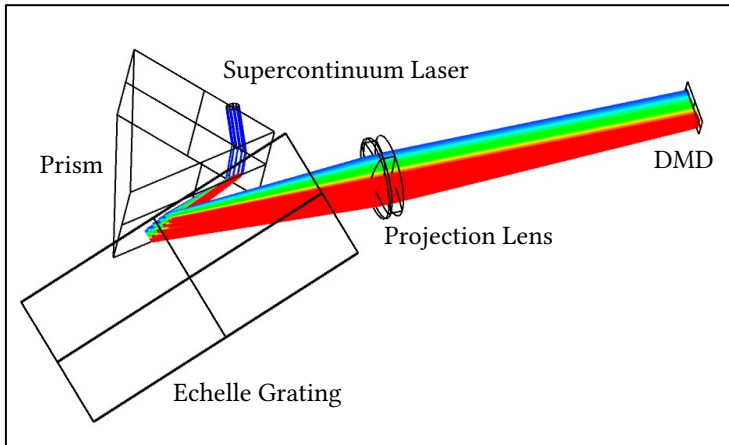


Figure 3.3: System schematic of the echellogram system in the sequential mode of OpticStudio.

To demonstrate the orthogonal dispersion generated by the two dispersers, Figure 3.4 presents the footprint diagrams at two apertures: the echelle grating surface (left) and the first surface of the achromatic doublet after the echelle grating (right). The light beams with wavelengths of 468.1 nm and 472.1 nm at the 119th order and 697.6 nm and 706.6 nm at the 78th order are illustrated and colored by their respective wavelength. As can be seen on the left, after the prism and before the echelle grating, light is only dispersed in the horizontal direction. After the echelle grating, vertical dispersion is introduced as shown by the diagram on the right.

A program is written in Python with PyZDDE [Sin15] package to automate the computation of diffraction orders and the setting of multi-configuration. Totally 42 configurations are utilized, each representing an order in the range from the 78th order to the 119th order. Within each order, 21 wavelengths are specified, equidistantly covering the free spectral range in that order. The system is telecentric and it has been optimized in terms of the spot size, which is contained within the Airy disk, offering diffraction limited optical quality. As an example, Figure 3.5 illustrates the spot diagram of 558.6 nm at the 99th diffraction order.

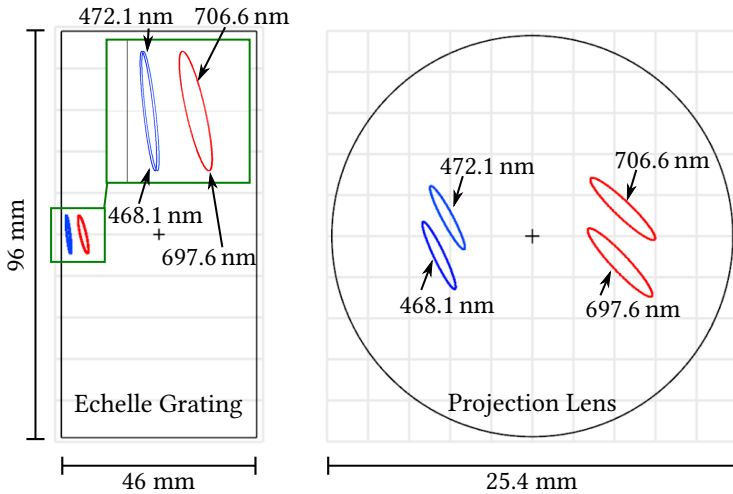


Figure 3.4: Footprint diagrams of the 119th order (blue) and the 78th order (red). Left: before the echelle grating surface. The green square shows a zoomed-in illustration. Right: after the echelle grating and before the achromatic doublet.

To characterize the spectral resolution of each DMD pixel, the diffraction ensquared energy fraction is calculated for a specific position while varying the wavelength to generate the spectral response of the underlying pixel. The pixel is placed at the centroid of the focused spot at 558.6 nm. The pupil sampling resolution is specified as 512×512 and a wavelength range from 558.5 nm to 558.7 nm is investigated. The goal is to characterise the bandwidth of light falling on this particular pixel. As illustrated by Figure 3.6, the spectral response of one pixel with width of $7.6 \mu\text{m}$ has an FWHM of 0.065 nm. Although this value varies with respect to the wavelength, the order of magnitude remains the same. When a larger area of 10×10 pixels is investigated, the resulting spectral width is slightly increased to 0.108 nm. As the size of the airy disk is relatively large with respect to the DMD pixel, diffraction effect is dominating the calculation of the ensquared energy fraction. This leads to the non-linear increase of spectral width when the detector area is increased. Since these values represent the optimum situation in theory, the spectral

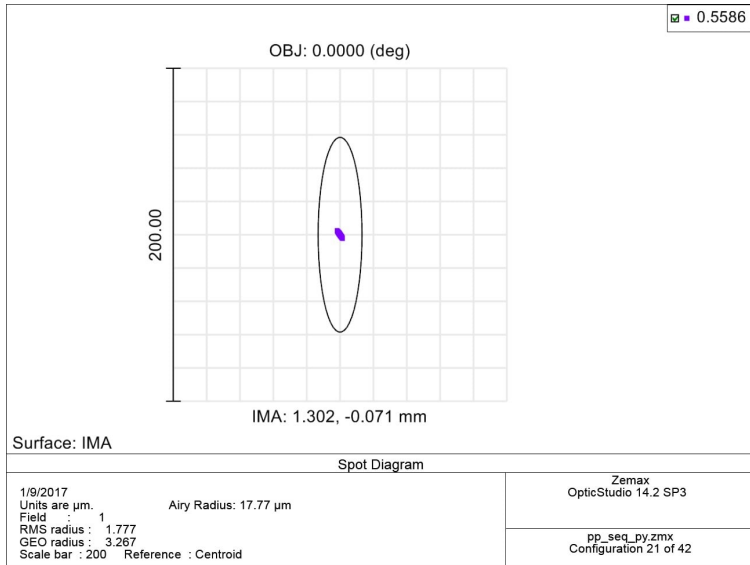


Figure 3.5: Spot diagram of 558.6 nm at the 99th diffraction order. The focusing spot on the DMD plane is shown in purple and the Airy disk is shown by the black ellipse in the graph.

widths will be further widened in practice due to imperfect alignment and tolerances of the optics.

3.1.1.3 Non-sequential Simulation

Once the optimization is complete, the system is reimplemented in the non-sequential mode of OpticStudio to simulate the generation of the echellogram and the effect of the DMD. Each pixel of the DMD has two stable states, namely “on” and “off”, where the micro mirror is turned by 12° and -12° respectively. Although the DMD in practice has 1920×1080 pixels, due to internal speed limitations of OpticStudio, the DMD used in simulation is set to 192×108 pixels. As illustrated in Figure 3.7, several additional components are placed compared to the echellogram in the sequential mode. When the pixels of the DMD are turned on, a coupling lens is used to focus the reflected light into

a liquid light guide. When the pixels of the DMD are turned off, the light is deflected into a beam trap. Therefore, by sending a specific pattern to the DMD, certain wavelength components can be selected and coupled into the light guide for further applications.

Distribution of the light on the DMD is simulated through ray tracing in the non-sequential mode. Totally 400 wavelengths are specified for each diffraction order and are divided into 20 groups. The spectral intensities of all wavelengths are assumed to be equal. Ray tracing is conducted for each group with 40000 rays. A detector color object with 1920×1080 pixels is placed right above the DMD to collect the projected rays. Another Python program is written to automate the switching between groups and orders. Figure 3.8 illustrates the simulation result.

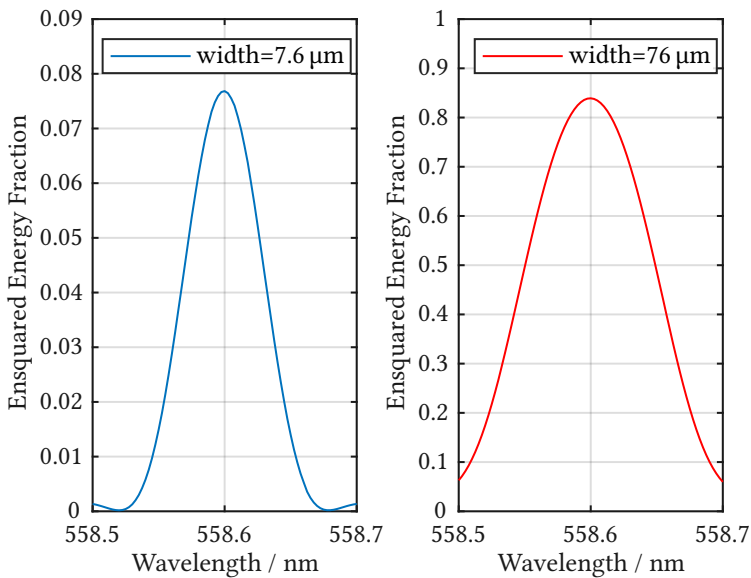


Figure 3.6: Spectral resolution characterized by the diffraction ensquared energy. Left: calculated with one pixel. Right: calculated with a block size of 10×10 pixels.

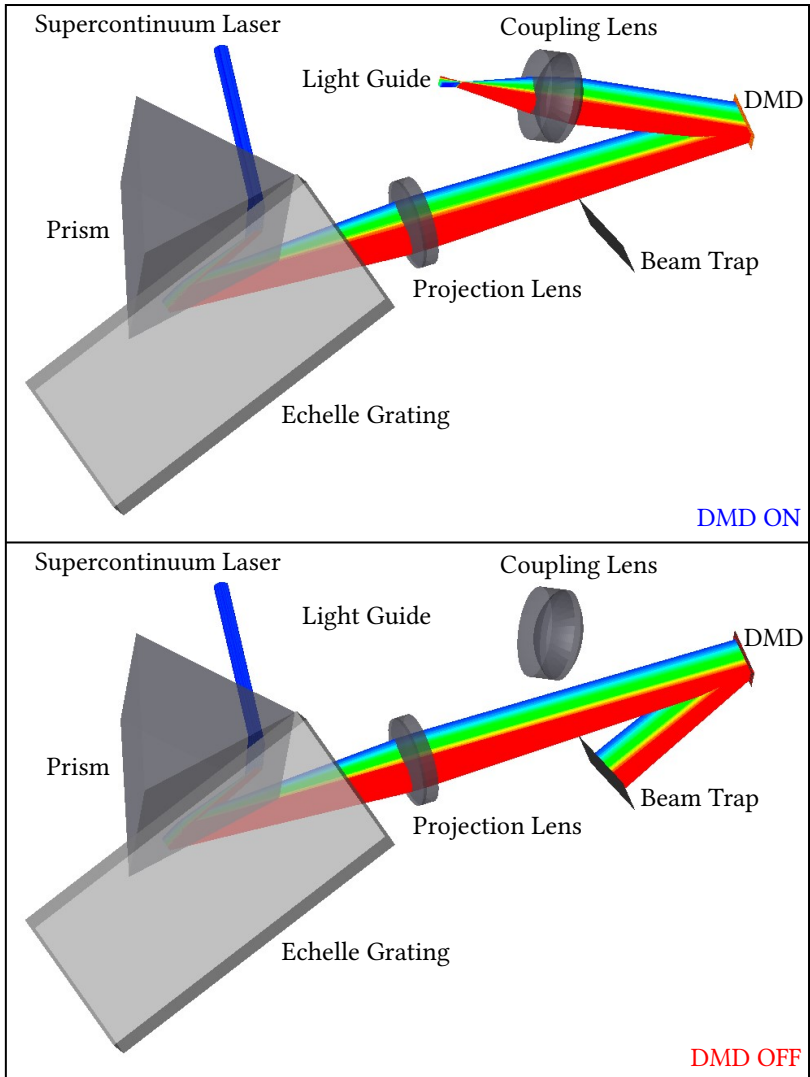


Figure 3.7: System schematic of the programmable light source in the non-sequential mode of OpticStudio. Upper: all pixels are on. Lower: all pixels are off.

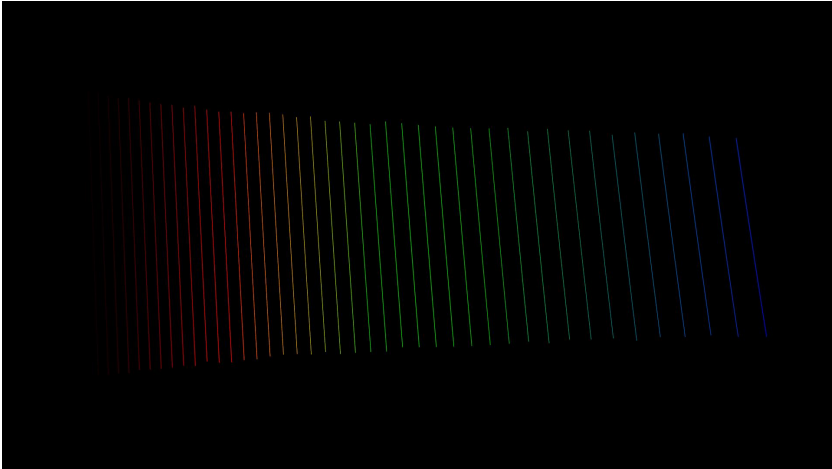


Figure 3.8: Simulated echellogram upon the DMD chip based on non-sequential ray tracing.

A detector is placed at the focus of the coupling lens to investigate the intensity distribution at the entrance of the light guide. As shown in Figure 3.9, when all pixels of the DMD are turned on, the reflected light forms an irregular spot through the coupling lens onto the detector. In practice, a light guide with a diameter of 5 mm is utilized to collect as much light as possible, which is shown as the red circle in Figure 3.9.

Two things should be noted regarding the design and simulation implemented in OpticStudio. Firstly, the grating efficiency is not taken into consideration during the simulation. For each order, the energy is assumed to be distributed evenly within the free spectral range. In practice, according to characteristics of the echelle grating, most of the energy will be concentrated around the blazing angle, i.e., within the free spectral range of each order. Nevertheless, a small portion of the energy will also spread to other directions, partly due to the imperfect blazing facet structure. Secondly, only geometric ray tracing is conducted when investigating interaction between light and DMD. In reality, considering the small size of the DMD pixel, the diffraction effect cannot be ignored [Woo12]. With its periodic structure, the DMD acts like a blazed

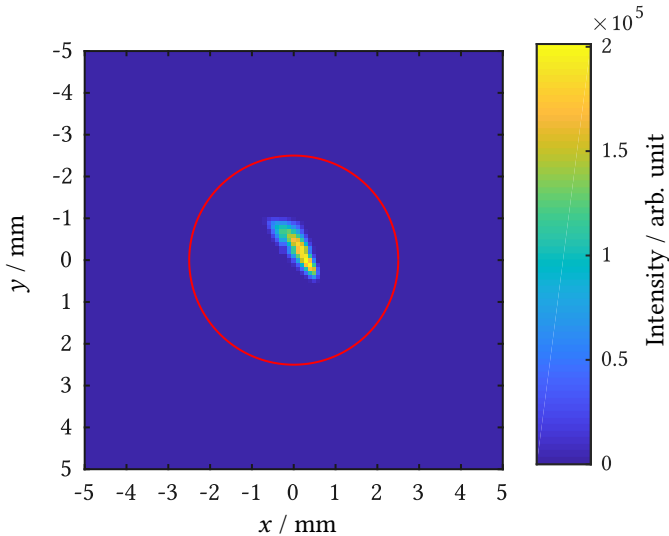


Figure 3.9: Irradiance distribution at the entrance of the liquid light guide. The red circle indicates the aperture of the light guide which has a diameter of 5 mm.

grating with a switchable blazing angle, which generates multiple 2D diffraction orders instead of simple reflection. Therefore, simulation of the intensity distribution in Figure 3.9 is only a rough approximation of the real scenario.

3.1.2 Setup and Alignment

The primary light source in the system is an obsolete model of supercontinuum laser from Koheras (now NKT Photonics). It is similar to the SuperK EXTREME EXW-12 model from NKT Photonics, which delivers 1.2 W of power in the range from 350 nm to 850 nm. As shown in Figure 3.10, the collimated output of the supercontinuum laser is first expanded with a $4\times$ reflective beam expander (BE04R/M from Thorlabs). The expanded beam then gets reflected by a visible mirror so that the infrared component passes through the mirror and enters a beam trap. The reflected beam passes through the prism and gets dispersed in the horizontal direction. As mentioned in the previous section,

the echelle grating has a groove density of $31.6 \text{ grooves mm}^{-1}$ and a blazing angle of 63.5° . The grating is manufactured by Richardson Grating Lab using Zerodur for substrate and aluminum for coating. After vertical dispersion generated by the echelle grating, the two-dimensionally dispersed light is focused onto the DMD with an achromatic doublet (AC254-100 from Thorlabs). The DMD used in the system is DLP LightCrafter 6500 EVM from Texas Instrument. The DMD chip has 1920×1080 pixels with a pitch of $7.6 \mu\text{m}$. Light of wavelengths corresponding to pixels that are turned on is collected by a second achromatic doublet with a shorter focal length (AC254-30 from Thorlabs) and coupled into a liquid light guide with a diameter of 5 mm. A second beam trap is placed in the opposite position to collect light with wavelengths corresponding to pixels that are turned off.

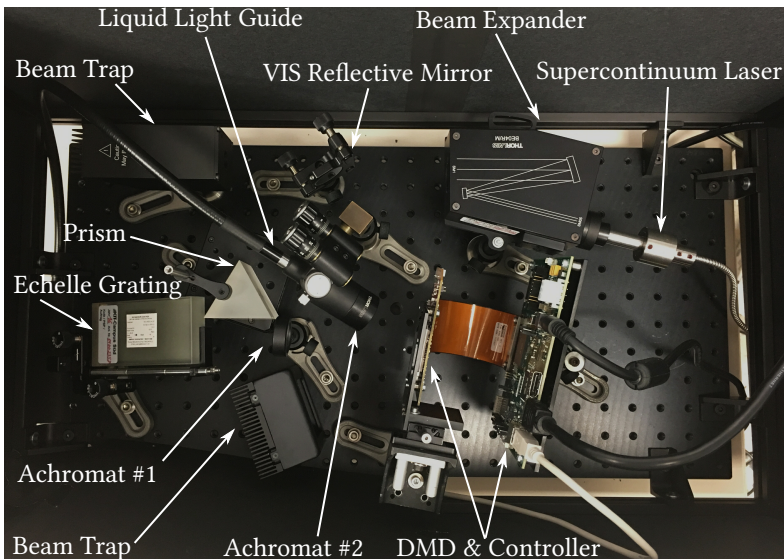


Figure 3.10: System setup of the programmable light source within its encapsulation.

Although the system is composed of relatively few components, the alignment proves to be not trivial. To begin with, laser safety is a major issue throughout the alignment process. As a Class 4 laser, the supercontinuum laser utilized

in the system is not eye safe even when operated at 1% of power. At earlier stage of calibration as well as under scenarios where higher power is required, an augmented reality setup based on Oculus Rift and a webcam is utilized. A Python program with the OpenCV package [Bra00] feeds image from the webcam to the Oculus Rift with proper distortion correction, in order to avoid any contact between the eyes of the operator and the laser. In other situations, a laser goggle with IRD5 filter from NoIR LaserShields is used and an additional neutral density (ND) filter is applied to the laser. Once alignment is finished, the complete system is encapsulated with a cage made of anodized aluminum rails and black cardboard, so that the calibration can be made without further laser safety protection. In general, alignment with a supercontinuum laser is always difficult, since a proper laser goggle with sufficient filtering will render the environment and non-illuminating components too dark.

Secondly, as shown in Figure 3.8, the positions of the prism and the echelle grating are designed so that usage of the DMD area is optimized. In practice, taking the grating efficiency into consideration, the incidence angle at the entrance of the prism has to be modified together with all following components to find the optimum. What adds to the complexity of alignment is that the spectrum varies with the power of the supercontinuum laser. In particular, shorter wavelengths gain more power as the total power is increased, due to the nonlinear effect of supercontinuum generation. As more subtle adjustments are made using laser goggles as protection, which requires low power operation, the DMD space for shorter wavelengths can only be estimated and reserved. Since the complete measurement of the echellogram is only possible at the calibration stage, the circle of alignment and calibration has to be repeated several times before an optimum situation is found.

Thirdly, although the prism and the grating are designed to be placed side by side, as shown in Figure 3.7, the edge of the ruling area of the echelle grating exhibits lesser quality than the center area in practice, which may lead to deterioration of the SNR in the final system. Therefore, the echelle grating is moved to the left of the prism to allow for the usage of center area. Although angular positions of the orders remain the same, the spatial positions are more separated at the projection lens, which adds more spherical aberration to the

edge orders. Such aberration is considered acceptable since the spot size is well below diffraction limit, as shown in Figure 3.5.

Last but not least, the DMD exhibits very strong diffraction effects, i.e. the reflection is composed of multiple two-dimensional orders. The rotation of the DMD and the position of the coupling lens has to be manipulated iteratively to have maximum collected power while maintaining acceptable spot size upon the DMD.

3.1.3 Spectrum Generation

The spectrum of the programmable light source is controlled with the pattern displayed on the DMD, which reflects the desired part of the wavelengths into the liquid light guide for output. The DMD is only capable of displaying binary patterns, since each micro mirror has only two stable tilting angles, which correspond to the state of on and off. Therefore, the intensity of each wavelength has to be manipulated through time multiplexing, i.e., multiple DMD patterns are combined to generate a spectral energy distribution within a certain period of time. Such a process is analyzed and discussed below.

To begin with, the temporal spectral flux of the programmable light source Φ_a can be expressed as

$$\Phi_a(\lambda, t) = \iint_{\mathbb{D}_a} r_a(x_a, y_a, t) E_a(x_a, y_a, \lambda) dx_a dy_a, \quad (3.4)$$

where r_a represents the temporal reflectivity of the spectral DMD and E_a denotes the spectral flux density. The lateral domain of the spectral DMD is represented by \mathbb{D}_a and the lateral coordinates of the spectral DMD are denoted by x_a and y_a .

For a certain duration of time, in which the reflectivity of the DMD changes, the average reflectivity of the spectral DMD can be defined as

$$R_a(x_a, y_a) = \frac{1}{T_a} \int_0^{T_a} r_a(x_a, y_a, t) dt, \quad (3.5)$$

where T_a denotes a certain duration of time.

Based on Equations 3.4 and 3.5, the spectral energy distribution of the generated light within a duration of T_a can be expressed as

$$\begin{aligned} Q(\lambda) &= \int_0^{T_a} \Phi(\lambda, t) dt \\ &= \int_0^{T_a} \iint_{D_a} r_a(x_a, y_a, t) E_a(x_a, y_a, \lambda) dx_a dy_a dt \\ &= \iint_{D_a} \int_0^{T_a} r_a(x_a, y_a, t) dt E_a(x_a, y_a, \lambda) dx_a dy_a \\ &= T_a \iint_{D_a} R_a(x_a, y_a) E_a(x_a, y_a, \lambda) dx_a dy_a. \end{aligned} \quad (3.6)$$

Since the DMD is composed of discrete micro mirrors in practice, Equation 3.6 can be rewritten in a matrix form:

$$\mathbf{q} = \mathbf{E}_a \mathbf{r}_a, \quad (3.7)$$

where $\mathbf{q} \in \mathbb{R}^{n_\lambda}$ represents the generated spectrum and $\mathbf{r}_a \in \mathbb{R}^{n_a}$ represents the required average reflectivity of the DMD pixels, which is reshaped into a 1D vector. The matrix $\mathbf{E}_a \in \mathbb{R}^{n_\lambda \times n_a}$ contains the spectral responses of all DMD pixels: $\mathbf{E}_a = [\mathbf{e}_1 \quad \mathbf{e}_2 \quad \dots \quad \mathbf{e}_{n_a}]$, where \mathbf{e}_i represents the spectral flux of the

i -th pixel. For a given spectrum, the required reflectivity of the DMD can be calculated by solving the following nonnegative least square problem:

$$\begin{aligned} & \underset{\mathbf{r}_a}{\text{minimize}} && \|\mathbf{E}_a \mathbf{r}_a - \mathbf{q}\|_2 \\ & \text{subject to} && \mathbf{r}_a \geq \mathbf{0}, \end{aligned} \tag{3.8}$$

where $\|\cdot\|_2$ represents the Euclidean norm.

Nevertheless, solving the problem with optimization algorithms can be very time consuming, especially considering the size of the calibration matrix. Therefore, the condition of nonnegativity is ignored and the problem is approximated by the following:

$$\begin{aligned} & \underset{\mathbf{r}_a}{\text{minimize}} && \|\mathbf{r}_a\|_2 \\ & \text{subject to} && \mathbf{q} = \mathbf{E}_a \mathbf{r}_a, \end{aligned} \tag{3.9}$$

which attempts to find the the least square solution with minimized norm. The solution can be easily calculated by applying the pseudo-inverse matrix of \mathbf{E}_a to both sides of the linear equation:

$$\mathbf{r}_a^* = \mathbf{E}_a^+ \mathbf{q} = \mathbf{E}_a^T (\mathbf{E}_a \mathbf{E}_a^T)^{-1} \mathbf{q}. \tag{3.10}$$

Since the FWHM of the fitted Gaussian in the calibration matrix is very narrow, the row rank of the calibration matrix is very close to full. Therefore, the solution by applying the pseudo-inverse matrix tends to be nonnegative in most of the cases. In rare circumstances where the reflectivity has negative values, the negative values are clipped to zero. As the pseudo-inverse matrix can be calculated off-line and pattern generation requires only one matrix multiplication, this method is much faster than solving the nonnegative least square problem rigorously with optimization and still provides acceptable results.

3.1.4 Calibration and Results

The calibration process aims to characterize the spectral response of each pixel so that any arbitrary spectrum can be generated by calculating the corresponding DMD pattern. The target is to generate the matrix E_a presented in the previous section.

3.1.4.1 Calibration Method

As shown in Figure 3.11, an additional setup is built outside of the encapsulated programmable light source. The liquid light guide is routed out of the encapsulation for calibration. Output light from this end of the liquid light guide is first collimated through an aspheric condenser lens (ACL5040U from Thorlabs), and then passes through a pair of microlens arrays (#63-230 from Edmund Optics) for homogenization. Lastly, an achromatic doublet is used to project the light to a rectangular area, where a fiber is placed which leads to a spectrometer. A special mounting adapter is machined in-house to hold all components at the correct positions. The combination of the condenser, the microlens arrays and the achromat is selected to match the diameter and the numerical aperture of the liquid light guide output, so that most of the energy is uniformly concentrated in a central rectangular block of 16 mm \times 12 mm, with minor portion of energy leaked to adjacent blocks. The homogenized rectangular illumination can be directly used in various applications once the calibration is finished.

To make the calibration, firstly all DMD pixels are turned on. The position of the fiber end is aligned so that the overall intensity of the spectrum is the highest. Then each pixel of the DMD is turned on while the corresponding spectrum is recorded. Due to the limitation of the DMD speed, the signal-to-noise ratio of the spectrometer and the computer storage, instead of scanning all pixels individually, square blocks of pixels are grouped together to form macro pixels, which are scanned and measured in practice. The size of the block is chosen to be 5 \times 5 pixels, in order to balance the calibration speed with the resolution. The integration time of the spectrometer is set at 5 ms.

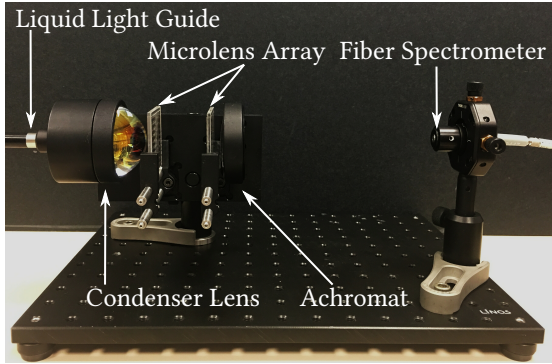


Figure 3.11: Calibration setup with homogenized rectangular illumination area projected onto a fiber leading to the spectrometer.

And the average intensity from five measurements is recorded for each macro pixel. The spectrometer used for calibration is the HR2000+ model from Ocean Optics, which covers the range from 190 nm to 1100 nm with a resolution of 0.66 nm and a step size of 0.44 nm.

Once the spectra corresponding to all macro pixels are measured, the wavelength range is cropped to reduce the computational effort and the intensities are assembled into a 2D array, where one axis represents the wavelength and the other axis represents the location of the macro pixel. Firstly, an intensity mask is generated by applying a 1D median filter along the axis of location, which is subtracted from the original array in order to remove fixed pattern noise of the spectrometer and intrinsic drift of the supercontinuum laser spectrum. The 2D array is then reshaped into a 3D hyperspectral cube, where two axes represent the position of the macro pixel on the DMD, and one axis represents the wavelength. Afterwards, 2D Wiener filtering is applied to each wavelength layer for adaptive noise removal. The filtered hyperspectral cube is denoted by E_{3D} and is reshaped back into a 2D array. After spectra with maximum intensity lower than the predefined threshold are discarded, the rest of the spectra are fitted to a Gaussian peak. All Gaussian peaks are combined into a 2D calibration matrix E_a , where the row index of E_a represents

the wavelength and the column index of E_a denotes position of the macro pixel.

3.1.4.2 Calibration Result

During the calibration procedure, a 3D hyperspectral tensor $E_{3D} \in \mathbb{R}^{n_\lambda \times n_x \times n_y}$ is generated containing measured spectra for all macro pixels of the DMD, where n_x and n_y represent the columns and rows of the macro pixel. By combining the wavelength layers in the tensor E_{3D} , the echellogram upon the DMD can be synthesized. As shown in Figure 3.12, the actual echellogram is rotated with respect to the simulated result shown in Figure 3.8 due to the rotation of the DMD coordinate system in the alignment process. Although most of the energy are concentrated within the free spectral range of each order, a small part of the energy gets leaked out of the free spectral range. Therefore, for certain wavelengths, the energy gets distributed into two/three different orders/locations. The arc stripe where intensity is slightly reduced in the lower part of the synthesized echellogram might be caused by the groove structure anomaly of the echelle grating. It should be noted that the resolution of the echellogram synthesis is limited by the spectral resolution of the spectrometer. To be more specific, a spectrometer with better spectral resolution is capable of generating sharper synthesized echellogram. In general, the prism is able to separate the orders well enough as expected and the area of the DMD is fully utilized.

Although the tensor E_{3D} can be reshaped to directly generate the target matrix E_a by flattening the two lateral dimensions, in the practical calibration procedure, the measured spectral peak of each effective macro pixel is fitted to a Gaussian function to reduced the noise in the signal. Macro pixels with a maximum intensity smaller than a threshold are discarded due to low SNR. In Figure 3.13, the relationship between the FWHM and the peak wavelength of the fitted Gaussian is illustrated by plotting the fitting result of each macro pixel as a dot. The average FWHM of the fitted Gaussians is 1.02 nm, which indicates that the calibration resolution is very likely limited by the spectral resolution of the spectrometer (approx. 0.66 nm) as well as its pitch size (0.44 nm).

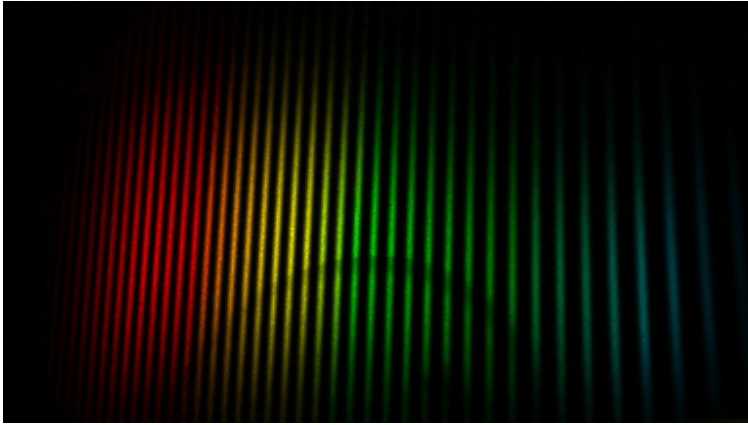


Figure 3.12: Synthesized echelogram upon the DMD chip generated from spectral measurements of the scanning macro pixels.

Therefore, the actual peak width of the spectrum corresponding to one macro pixel is potentially smaller than the currently measured value.

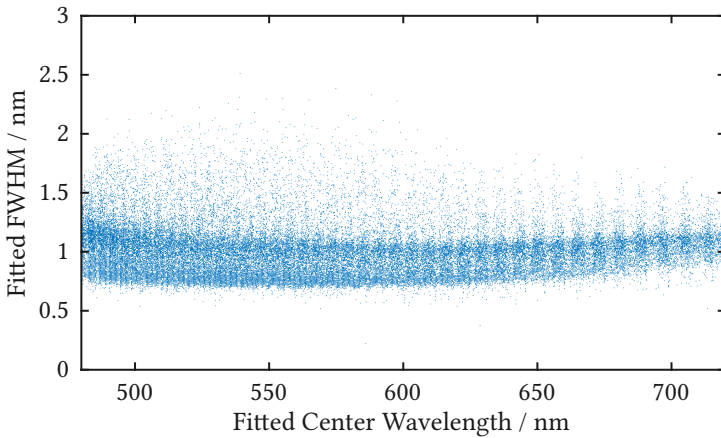


Figure 3.13: FWHM with respect to center wavelength of the fitted Gaussian peaks. Measurements corresponding to 54338 macro pixels are drawn.

The amplitude of the fitted Gaussian peaks are also plotted in Figure 3.14. Certain periodic variation of the maximum amplitude can be observed from the figure. This is due to the fact that not all energy is concentrated within the free spectral range of each order. For wavelength at the center of the free spectral range of one diffraction order, most energy is concentrated in the corresponding macro pixel, thus achieving higher maximum amplitude for the fitted Gaussian function. For wavelengths at the edge of the free spectral range, part of the energy will be distributed to the adjacent order, making the maximum amplitude for this wavelength lower.

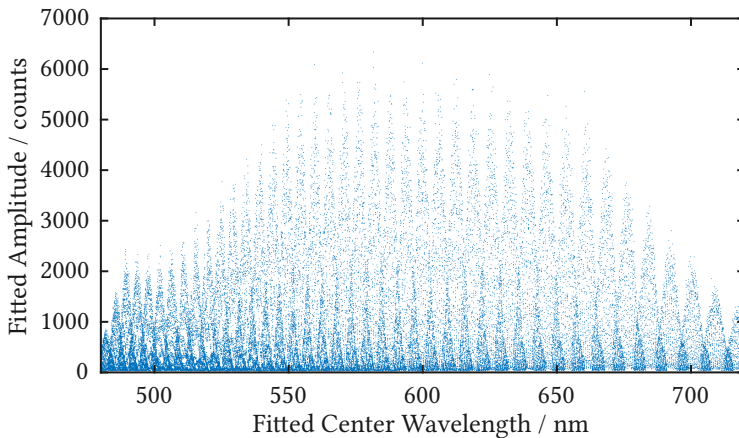


Figure 3.14: Amplitude with respect to the center wavelength of the fitted Gaussian peaks. Measurements corresponding to 54338 macro pixels are drawn. The amplitude is shown in count, which is the native unit of the spectrometer.

To evaluate the effect of the calibration, especially the fitting process, all fitted Gaussian signals of the effective macro pixels are summed and compared against the total spectrum measured when all pixels are turned on. The total spectrum is measured using a much smaller integration time in order to avoid saturation of the spectrometer, and is therefore scaled according to the ratio of integration times. As can be seen in Figure 3.15, the sum of the selected macro pixels is quite close to the spectrum when all pixels are turned on. In fact, only 22% of the macro pixels are selected for calibration, which

contributes to the majority of the energy, whereas the rest of the pixels are discarded for their poor signal-to-noise ratio.

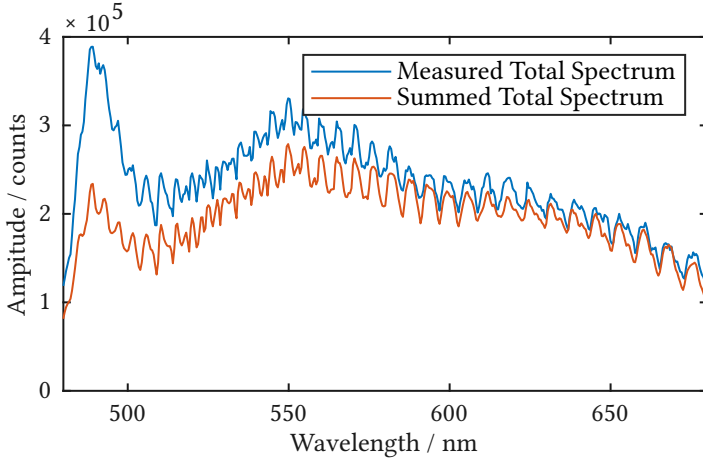


Figure 3.15: Comparison between measured total spectrum and calculated spectrum by summing all useful macro pixels.

Once the average reflectivity is calculated based on Equation 3.10, it is re-shaped into the complete DMD size and quantized into an 8-Bit pattern, which can be displayed through time multiplexing. DMD patterns for several target spectra are calculated and the generated spectra are measured with an integration time of 5 ms.

Figure 3.16 illustrates the generation of a flat spectrum spanning across the range from 480 nm to 700 nm. Although with some variation of the intensity due to intrinsic spectral variation of the supercontinuum laser and noise, the generated spectrum is very close to the target. The overshoot on the edges of the flat spectrum can be explained by the Gibbs phenomenon, as the maximum frequency that the system is capable of generating is dependent on the FWHM of the fitted Gaussian for each macro pixel.

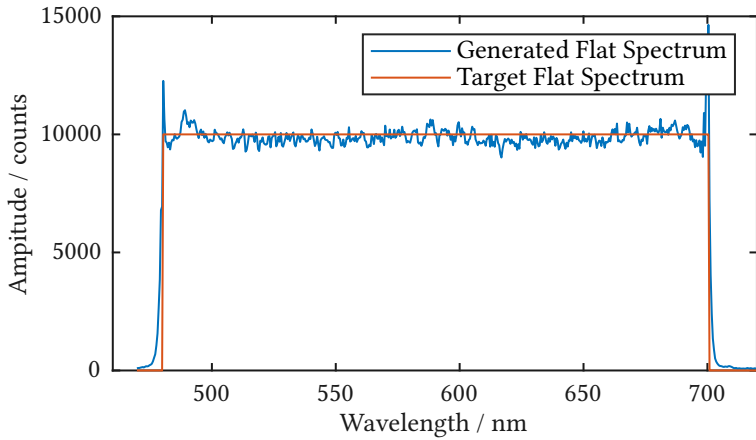


Figure 3.16: Comparison between measured flat spectrum and target spectrum.

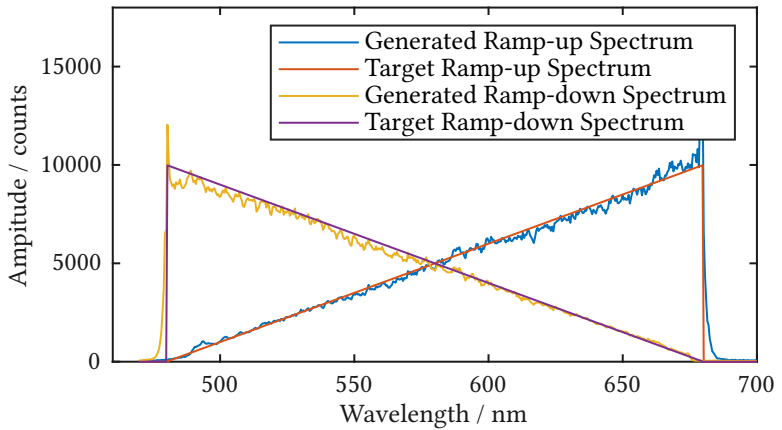


Figure 3.17: Comparison between measured ramp spectra and target spectra.

Figure 3.17 demonstrates two ramp spectra from 480 nm to 680 nm. Similar to the flat spectrum, overshoot can be observed during abrupt change at the edges while the rest of the spectra closely reproduce the targets.

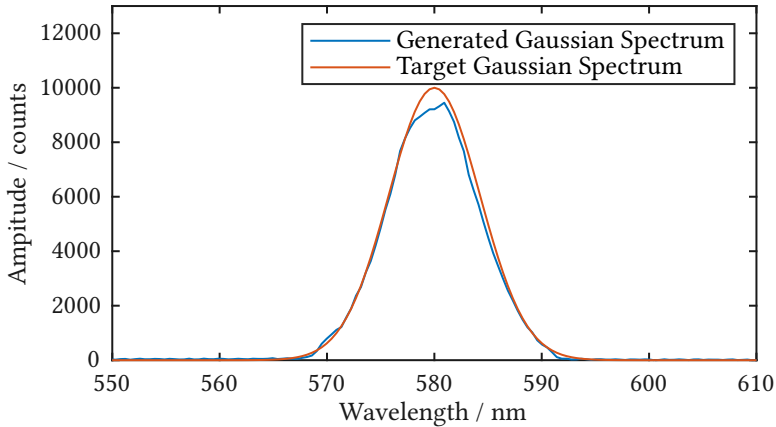


Figure 3.18: Comparison between measured Gaussian spectrum and target Gaussian spectrum.

Last but not least, Figure 3.18 shows a Gaussian spectrum centered at 580 nm with a FWHM of 10 nm. Gaussian peaks with narrower FWHM can be generated as well and the smallest possible FWHM is limited by the FWHM of the spectrum generated by a single macro pixel. As mentioned previously, the average FWHM of the fitted Gaussian for each macro pixel is 1.02 nm, which is very likely limited by the resolution of the spectrometer in the current calibration setup.

3.1.4.3 Wavelength Tuning Resolution

As the spectrum of the system can be programmed arbitrarily, the combination of the echelle grating and the DMD can also be treated as a tunable band pass filter, achieving a function similar to an acousto-optic tunable filter (AOTF) which is often applied to the supercontinuum laser to realize wavelength scanning.

To characterize the wavelength tuning resolution of the current system, a series of Gaussian spectra are generated, each having a FWHM of 2 nm. The center wavelength of these spectra ranges from 560 nm to 580 nm with a step size of 0.01 nm. The calibration setup mentioned previously is used to record

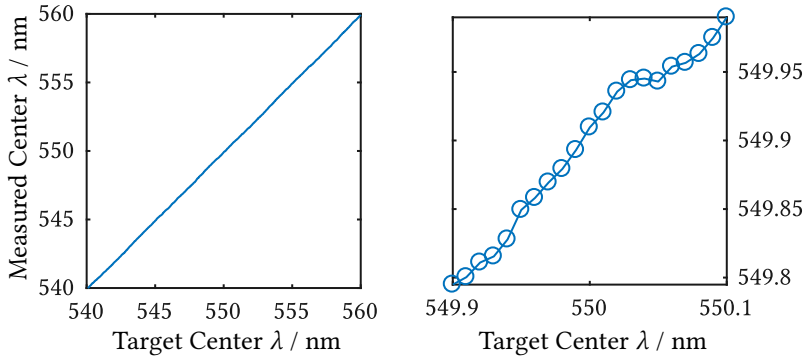


Figure 3.19: Scanning of Gaussian spectra with FWHM of 2 nm. Left: wavelength range from 540 nm to 560 nm. Right: zoomed wavelength range from 549.9 nm to 550.1 nm.

the generated spectra with an integration time of 5 ms. Gaussian fitting is then applied to all recorded spectra to yield measured center wavelengths, which are compared against the target. As shown in Figure 3.19, the measured scanning results demonstrates superior linearity with very few errors. When zoomed into a smaller region of 0.2 nm, it can be seen that a step size of 0.01 nm can be correctly reflected by the measurement with minor errors caused mainly by noise. As an example, Figure 3.20 illustrates five measured spectra from the scanning sequence.

3.1.5 Discussion

In this section, a novel programmable light source in visible range is presented, which utilizes a supercontinuum laser as the primary source and combines its echellogram with digital mirror device for programmable spectral filtering. The echellogram is firstly designed in the sequential mode of OpticStudio to have diffraction limited telecentric imaging upon the DMD. Then the complete system, including the DMD and the coupling of the output light into a liquid light guide, is simulated in the non-sequential mode with ray tracing to generate an echellogram image showing the free spectral range of diffraction orders No.78 to No.119.

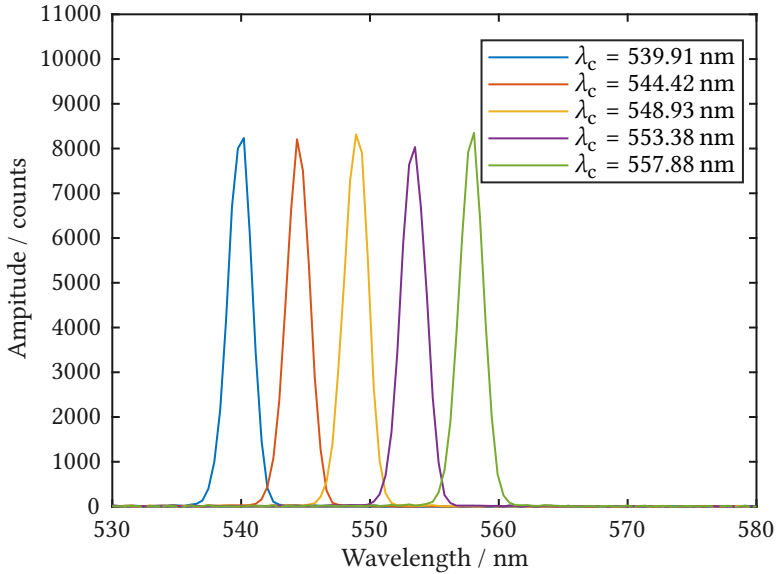


Figure 3.20: Five Gaussian spectra from the scanning sequence. The fitted center wavelength is represented by λ_c .

The system is constructed in the lab and evaluated. To calibrate the programmable filter, the output of the liquid light guide is homogenized and projected onto a fiber spectrometer. Pixel blocks of 5×5 are scanned while their corresponding spectral responses are recorded. The recorded data are cleaned, smoothed and fitted with several processes before the calibration matrix is constructed by assembling fitted Gaussian spectra of the useful macro pixels. Average FWHM of the fitted Gaussians is 1.02 nm, which is believed to be limited by the resolution and the step size of the spectrometer. During the calibration procedure, intensity measurements of different wavelengths can be combined to synthesize the echellogram upon the DMD. Two major differences are observed between measured echellogram and the simulated result from ray tracing. Firstly, small part of the energy gets leaked out of the free spectral range of each order due to imperfect blazing structure of the echelle grating dependent on the manufacturing accuracy. Secondly, width of

each order is widened compare to the simulated results, since the ray tracing simulation does not take into account the diffraction limit, as is illustrated in Figure 3.5.

Several exemplary spectra are generated and compared against the target, which demonstrates that the spectral filtering is relatively accurate. A series of Gaussian spectra are generated and measured to investigate the wavelength tuning resolution of the system when operated as a scanning source. Results have shown that the system is responsive to a step size of 0.01 nm.

Currently three major factors are limiting the performance of the system. Firstly, the supercontinuum laser, which is used in the system as the primary input to the programmable spectral filtering system, has a limited spectral stability. On one hand, the intrinsic spectral variation of the supercontinuum laser gets directly passed to the final output. On the other hand, as the calibration process takes a relatively long period of time, the spectral variation is also transferred into the calibration matrix, reducing its accuracy. Secondly, like most echellogram systems, the programmable filtering setup is very sensitive to mechanical vibrations, since tiny movement of the echelle grating shifts wavelengths by multiple pixels. Currently Sorbothane feet are attached to the breadboard to absorb vibration. Nevertheless, optical table with active self-leveling isolators would definitely enhance the stability of the system. Last but not least, resolution and step size of the USB fiber spectrometer is limiting the calibration accuracy. A spectrometer with less measurement range but higher resolution would be more suitable for the calibration task.

In summary, the proposed system is potentially useful for any optical applications where manipulation of the wavelength is necessary. In particular, the system provides a versatile prototyping platform for measurement systems based on chromatic principle as well as hyperspectral imaging technologies. The output light from this programmable light source is directly inserted into the programmable array microscope to form the AdaScope, based on which various measurement methods are developed.

3.2 Programmable Array Microscope

In this section, a programmable array microscope with axial chromatic encoding is presented. Using the aforementioned programmable light source as the illumination source, the presented system is capable of adaptively changing its measurement mode through electronic control of the light source as well as the programmable array.

3.2.1 System Design

The programmable array microscope setup is similar to the conventional reflective confocal scanning microscope, except that a DMD is used as a spatially-programmable light source and a camera is used to measure all lateral locations simultaneously. As a reflective setup, the system can be considered as being composed of two parts, i.e., the illumination arm and the imaging arm, which share the same chromatic optics for axial chromatic encoding.

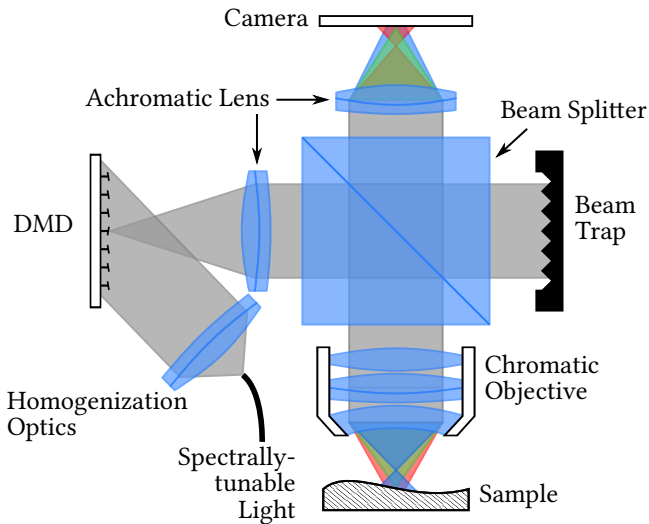


Figure 3.21: System schematic of the programmable array microscope with chromatic encoding.

As illustrated in Figure 3.21, light transported through the liquid light guide from the programmable source is first collimated and homogenized through a group of homogenization optics before the light is projected into a rectangular illumination field upon the DMD using an achromatic doublet. The DMD model used in the microscope setup is the same model as the one in the tunable light source. Each pixel of the DMD acts as a secondary point source which can be programmably addressed. After passing through a series of collimation lenses and getting reflected by the pellicle beamsplitter (BP245B1 from Thorlabs), light from selected DMD pixels is projected onto the target sample using an objective lens with designed chromatic separation along the optical axis (Precitec CLS4). The reflected light passes through the chromatic objective once again and gets focused onto an sCMOS camera (Andor Zyla 5.5).

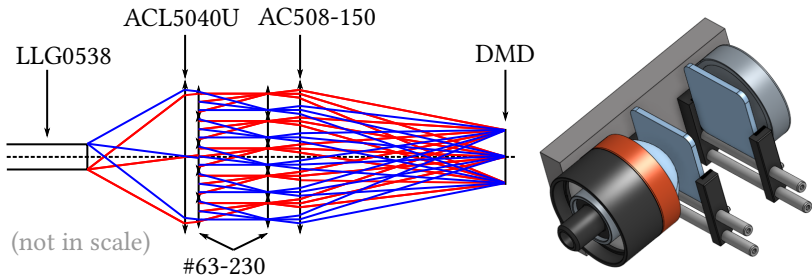


Figure 3.22: Schematic and CAD rendering of the beam homogenization optics.

The major difference between the programmable array microscope and a conventional single point scanning microscope is the additional requirement of homogenized illumination on the programmable pinhole array. In the AdaScope system, a beam homogenization module has been developed based on the combination of a condenser lens, two microlens arrays and an achromatic doublet, as shown in Figure 3.22. Two microlens arrays are utilized since both the diffraction effect and the flat-top broadening are minimized compared to a single microlens array [Bue02]. The specifications of the condenser lens, the microlens arrays and the achromatic doublet are specifically chosen in combination according to the output diameter and NA of the liquid light guide, so

that the resulting rectangular illumination field is only slightly larger than the effective area of the DMD, thus fully utilizing the power of the programmable light source. To get the optimum configuration of the various components, the exit facet of the liquid light guide is treated as an extended light source. Each point of the facet emits a bundle of light rays (shown as blue or red in Figure 3.22), which is collimated by a condenser lens. Depending on the position of the point source on the facet, the collimated light forms a certain angle with respect to the optical axis. Then the collimated light immediately passes through a microlens array generating a two-dimensional array of focal points. A secondary microlens array with the same specification is placed at the focal plane of the first array to rectify the beam bundle of each focal point. Lastly, an achromatic lens projects the light rays onto a rectangular field. A customized adapter has been designed and manufactured for the mounting of all components.

3.2.2 Simulation and Construction

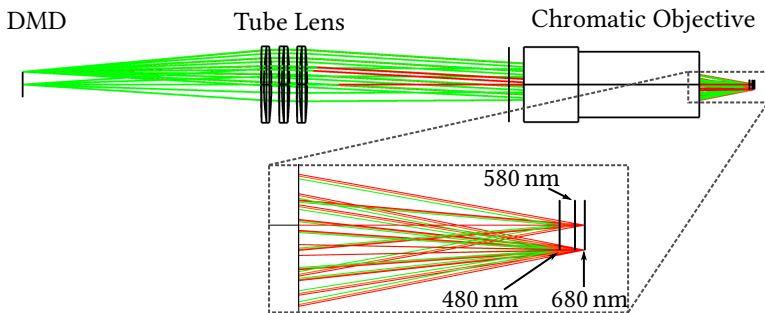


Figure 3.23: Simulation of the illumination arm.

The microscope system is fully simulated and analyzed with OpticStudio in the sequential mode. The wavelength range of the system is specified to be from 480 nm to 680 nm and a black-box model of the chromatic objective is adopted from the manufacturer. As demonstrated by Figure 3.23, a group of three achromatic doublets are chosen to form an infinity-corrected tube lens,

which matches the field of view (FoV) and aperture of the chromatic objective with the effective area of the DMD as well as the illumination NA.

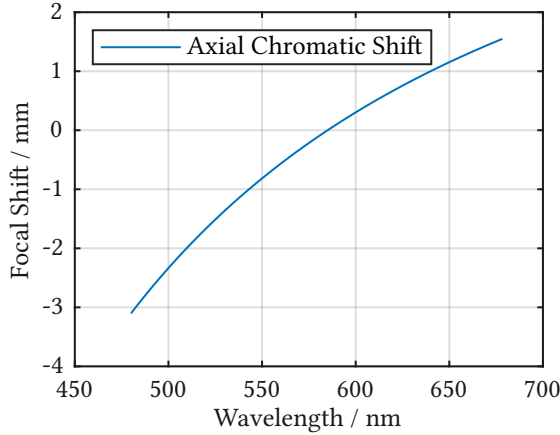


Figure 3.24: Axial chromatic shift in object space.

Due to the intrinsic design of the chromatic objective, the axial focus shift is slightly non-linear (Figure 3.24). Additionally, the field curvature has been slightly increased due to the introduction of the simple tube lens. Nevertheless, both factors can be easily corrected through experimental calibration and postprocessing. Apart from these aberrations, the designed combination of the tube lens and the chromatic objective generates diffraction limited illumination spots in the object space for all wavelengths at their corresponding focal plane. The imaging arm is the same as the illumination arm with an identical tube lens placed before the camera and the camera sensor is located at the conjugate location of the DMD. When a mirror is placed at the illumination focal plane of each wavelength, diffraction-limited focus points are achieved at the camera plane. For wavelengths of 480 nm, 580 nm and 680 nm, the airy disk has a diameter of 4.5 μm , 5.4 μm and 6.4 μm respectively, which matches the pixel size of 6.5 μm in the camera sensor. The three dimensional measurement volume is roughly 5.4 mm (x) by 3 mm (y) by 4.67 mm (z) for the specified wavelength range.

Figure 3.25 demonstrates the constructed programmable array microscope setup. The system is aligned through the following steps. Firstly, axial position of the camera is fixed so that perfect focus is achieved through the tube lens when a parallel light source is used from the direction of the beam trap. The lateral position of the camera is defined by the mechanical tube connection between the camera and the beamsplitter assembly. Secondly, axial positions of the DMD and a patterned reflective sample are adjusted iteratively until the intrinsic pattern from the sample and the DMD pattern are both in focus for illumination with a single wavelength. The lateral position of the DMD is adjusted in the meanwhile to remain as centered in the camera image as possible. Thirdly, angle of the DMD is adjusted so that perfect focus is achieved through the complete field. Last but not least, illumination angle of the beam homogenization setup is adjusted to meet two conditions. On one hand, captured intensity on the camera should be as high as possible for better SNR. On the other hand, blurring of a single focal spot is checked by shifting a reflective sample axially so that the blurred spot is as balanced and homogenized as possible. In practice, multiple iterations through these steps are performed to achieve an optimum alignment.

3.2.3 Camera Calibration

One key aspect which separates the area confocal scanning system from a conventional single point system is its requirement of calibration between the camera sensor and the light source. Considering the telecentric design of the system, the camera is calibrated for a single wavelength of 555 nm, by placing a mirror at its focal plane. The registration is made only between the camera coordinates and the DMD coordinates, while the registration toward the object/world coordinate system is not considered. Therefore, all measurement results can be laterally presented in the DMD coordinate system by first projecting the DMD coordinate system onto the camera frame and then making an interpolation, as demonstrated in Figure 3.26. More details regarding the calibration procedure can be found in Appendix A.1.

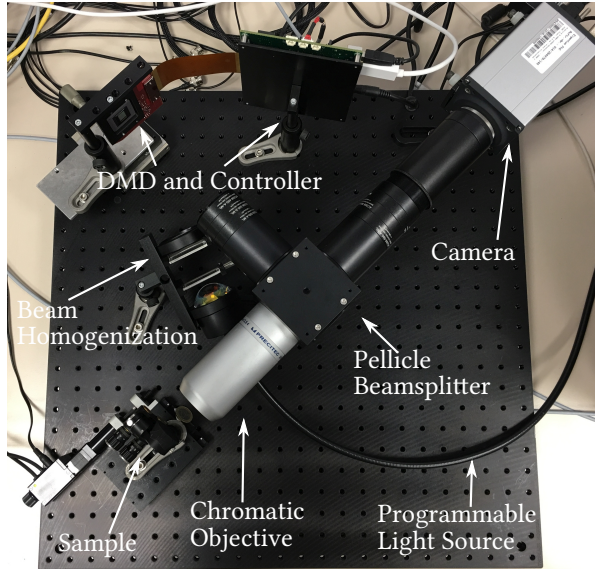


Figure 3.25: Setup of the programmable array microscope with chromatic encoding.

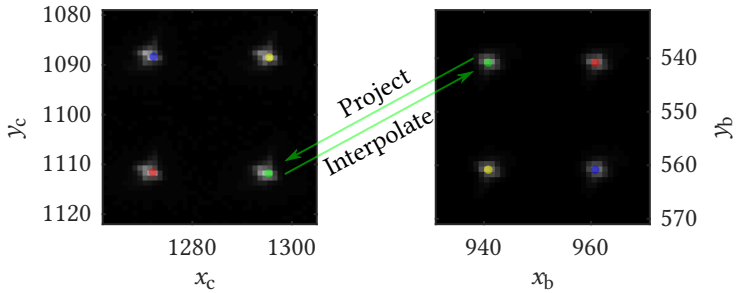


Figure 3.26: Camera calibration for Adascope. Left: camera coordinate system. Right: spatial DMD coordinate system.

With a broadband mirror as the target, the spectral sensitivity of the imaging system is also calibrated and incorporated into all measurement methods.

Variations in the spectral sensitivity mainly originate from the spectral response of the pellicle beam splitter as well as the camera sensor.

3.2.4 Illumination Generation

The adaptive nature of the AdaScope system mainly originates from its ability to generate arbitrary 3D illumination fields. The temporal spectral flux density over the spatial DMD can be calculated as

$$E_b(x_b, y_b, \lambda, t) = \frac{\Phi_a(\lambda, t)}{S_I}, \quad (3.11)$$

where S_I represents the area of the homogenized illumination from the programmable light source. The temporal illumination intensity distribution U can be computed through the 2D convolution between the temporal spectral flux density and a normalized chromatic intensity point spread function (PSF) $H_\lambda(x, y, z, \lambda)$. The lateral coordinates of the temporal spectral flux density have to be scaled by the paraxial magnification M_1 .

$$\begin{aligned} U(x, y, z, \lambda, t) &= c H_\lambda(x, y, z, \lambda) * \left[n_b \left(\frac{x}{M_1}, \frac{y}{M_1}, t \right) E_b \left(\frac{x}{M_1}, \frac{y}{M_1}, \lambda, t \right) \right] \\ &= \frac{c \Phi_a(\lambda, t)}{S_I} \iint_{-\infty}^{\infty} H_\lambda(x - x', y - y', z, \lambda) n_b \left(\frac{x'}{M_1}, \frac{y'}{M_1}, t \right) dx' dy' \end{aligned} \quad (3.12)$$

A constant factor represented by c is utilized to match the scaling and the unit. For a relatively sparse illumination pattern, crosstalk between adjacent illumination positions can be ignored and the intensity illumination function can be approximated by

$$H_\lambda(x, y, z, \lambda) \approx \delta(x) \delta(y) \delta(\lambda - g(z)), \quad (3.13)$$

where $g(z)$ represents the axial chromatic shift.

With this approximation, the illumination intensity distribution can be integrated over time and wavelength:

$$\begin{aligned}
 U(x, y, z) &= \int_0^{T_b} \int_{\mathbb{D}_\lambda} U(x, y, z, \lambda, t) d\lambda dt \\
 &= c \int_0^{T_b} \int_{\mathbb{D}_\lambda} \delta(\lambda - g(z)) \left[n_b \left(\frac{x}{M}, \frac{y}{M}, t \right) E_b \left(\frac{x}{M}, \frac{y}{M}, \lambda, t \right) \right] d\lambda dt \\
 &= c \int_0^{T_b} n_b \left(\frac{x}{M}, \frac{y}{M}, t \right) E_b \left(\frac{x}{M}, \frac{y}{M}, g(z), t \right) dt \tag{3.14} \\
 &= \frac{c}{S_I} \int_0^{T_b} \Phi_a(g(z), t) n_b \left(\frac{x}{M}, \frac{y}{M}, t \right) dt \\
 &= \frac{c}{S_I} \int_0^{T_b} n_b \left(\frac{x}{M}, \frac{y}{M}, t \right) \iint_{\mathbb{D}_a} r_a(x_a, y_a, t) E_a(x_a, y_a, g(z)) dx_a dy_a dt.
 \end{aligned}$$

The target of the illumination generation process is to find the combination of r_a and n_b which could generate the desired U with a minimum exposure time T_b while still maintaining all physical constraints. Similar to the process demonstrated in Equations 3.6 and 3.7, Equation 3.14 can also be discretized into a matrix form:

$$\mathbf{U} = \mathbf{E}_a \mathbf{R}_a \mathbf{R}_b^T, \tag{3.15}$$

where the matrix $\mathbf{U} \in \mathbb{R}^{n_z \times n_b}$ contains the 3D illumination intensities, which are reshaped into a 2D matrix. One axis denotes the n_z axial positions while

the other axis represents the n_b 2D lateral indices. The matrix $\mathbf{E}_{a'} \in \mathbb{R}^{n_z \times n_a}$ is generated based on \mathbf{E}_a through mapping between the wavelength and the axial position. The matrix $\mathbf{R}_a \in \mathbb{R}^{n_a \times n_t}$ contains n_t reflectance configurations of the n_a spectral DMD pixels. And the matrix $\mathbf{R}_b \in \mathbb{R}^{n_b \times n_t}$ contains n_t reflectance configurations of the n_b spatial DMD pixels. The pseudo-inverse of $\mathbf{E}_{a'}$ can be multiplied to both sides of the equation:

$$\mathbf{E}_{a'}^T (\mathbf{E}_{a'} \mathbf{E}_{a'}^T)^{-1} \mathbf{U} = \mathbf{R}_a \mathbf{R}_b^T. \quad (3.16)$$

The first step of finding the optimum configuration of the two DMDs would be to get a nonnegative full rank decomposition of the left side (which is not guaranteed to be nonnegative itself). This is not a simple task as the problem of nonnegative matrix factorization has been proven to be NP-hard [Vav10]. Additionally, the discretization of time steps is yet to be considered. Due to the bit-plane mixing capability of the DMD controller, the efficiency of multiple binary patterns is lower than, e.g., 8-bit patterns. Meanwhile, Equation 3.15 is equivalent to Equation 3.14 only when the discretization is based on binary patterns, i.e., at least one matrix from \mathbf{R}_a and \mathbf{R}_b has to be binary matrix. In practice, for most of the simple illumination distributions, either \mathbf{R}_a or \mathbf{R}_b is determined first in an empirical manner, while the other matrix is computed afterwards. Such a process is sufficient for the methods proposed in this thesis.

3.2.5 Synchronization Mechanism

The Zyla 5.5 sCMOS camera in the system has a very good signal-to-noise ratio (1.2 e^- read noise) and a high dynamic range (25000:1) but the speed (maximum 40 fps) is not particularly fast compared to some other high-speed camera models. To fully utilize its potential, the camera (controlled by the host computer) is applied as the master in the synchronization mechanism and triggers the spectral DMD in the programmable light source, which then triggers the spatial DMD in the area scanning microscope, as shown in Figure 3.27.

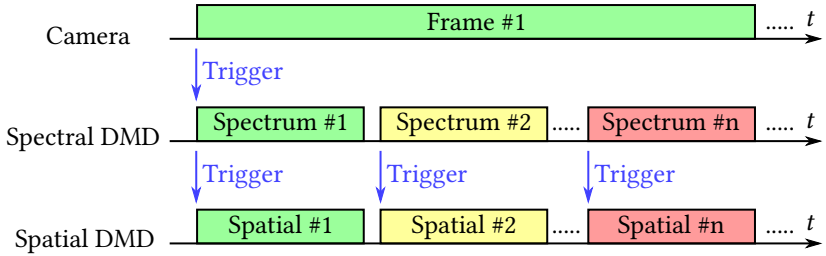


Figure 3.27: Triggering mode #1. The camera triggers the spectral DMD to start a series of illumination spectra. Each spectrum triggers a corresponding pattern on the spatial DMD.

Both DMDs are empowered with a binary pattern rate of 9.5 kHz, but the transmission of the patterns turns out to be the bottleneck of the pattern generation speed. Although the HDMI interface on the control board is capable of a wider transmission bandwidth, accurate synchronization of the two DMD becomes more difficult when both are connected through HDMI. Therefore, the on-board USB 1.1 interface is utilized in the pattern-on-the-fly mode, which allows for accurate triggering between the various components.

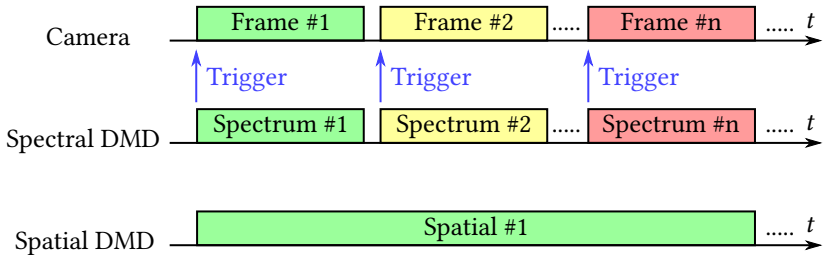


Figure 3.28: Triggering mode #2. The spectral DMD triggers the camera to capture a series of illumination spectra. The spatial DMD is operated directly through the host PC.

In certain situations, synced operation of the spatial DMD is not necessary so only synchronization between the camera and the spectral DMD is maintained while the spatial DMD is directly controlled by the host computer, as illustrated in Figure 3.28. In this case, it is particularly useful to apply the spectral DMD in video mode as the master which triggers the camera, so that

a larger number of complex patterns can be displayed by the spectral DMD without any interruption, which would have exceeded the internal memory of the DMD if operated in pattern-on-the-fly mode.

3.3 Summary

Based on a holistic design approach, the AdaScope system is composed of two parts. The programmable light source is developed based on the echellogram of a supercontinuum laser source, which is spatially filtered by a DMD. Through electronic control of the DMD, arbitrary output spectra can be generated in a time-multiplexed manner with high resolution and fast speed. The output light from the programmable light source is transmitted to a programmable array microscope through a liquid light guide, which is then homogenized and projected onto a secondary DMD to form an extended spatially programmable source. The microscope setup adopts a reflective configuration based on a beamsplitter and a chromatic microscopic objective is utilized for axial chromatic encoding.

In conventional confocal scanning systems, the scanning volume, which is formed by the scanning range of the three axes, can be seen as a three dimensional grid of discrete scanning locations, which are limited by the resolution of the scanning system. Each time only a fixed single point or point array can be measured. However, in the AdaScope system, through synchronization between the camera, the spectral DMD and the spatial DMD, an arbitrary combination of scanning locations within the scanning volume can be addressed within a single frame of exposure in a time multiplexed manner. Moreover, different scanning locations can be weighted in a single frame by manipulating the spectral intensity of the illumination light. As the key feature of the AdaScope, such adaptability allows novel measurement methods to be developed while the AdaScope is operated in different modes.

4 Cascade Measurement Strategy

Conventional 3D microscopic systems are dominated by the dilemma between scanning density and measurement accuracy, which is qualitatively expressed in Figure 4.1. Best accuracy and robustness is achieved when a single focal point is scanned in a confocal system (area A in Figure 4.1). As the density of simultaneously scanned locations increases, regardless of which measurement method is adopted, the degree of crosstalk also increases, which degrades measurement accuracy (area B in Figure 4.1). At a certain point, the confocal condition is no longer maintained and the system is reduced to a wide-field microscope (area C in Figure 4.1). Although methods based on shape from focus can be adopted, the measurement accuracy is generally worse than the equivalent confocal scanning system and is more sensitive to the underlying texture of the sample surface, which further degrades the robustness of the system.

Over the years, various measurement methods have been developed to push a particular section of the boundary toward the upper right direction. Instead of only relying on incremental improvements of any particular methods, the AdaScope system attempts to tackle this problem with a new kind of measurement strategy. Thanks to the intrinsic adaptability granted by the design and construction of the AdaScope, the system is capable of swiftly switching between different operational modes, i.e., between section A, B and C in Figure 4.1. For a complete measurement task, a cascade of measurement methods can be developed and applied, where raw and fast measurement result in one stage can be fed to the next stage of measurement as prior knowledge for more accurate measurements, as demonstrated in Figure 4.2. Such prior knowledge can be utilized by facilitating the initialization of the new measurement or

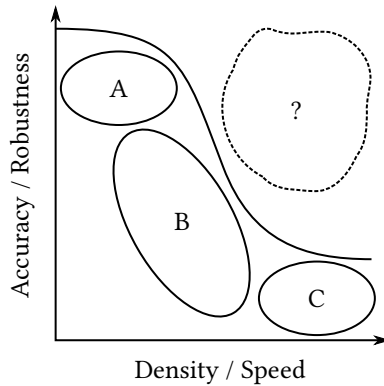


Figure 4.1: Dilemma of scanning density and measurement accuracy. A: single point confocal scanning. B: slit/array scanning. C: Shape from focus.

constraining the measurement range. This strategy allows advantages of each mode to be maximized, achieving optimum efficiency of the hardware.

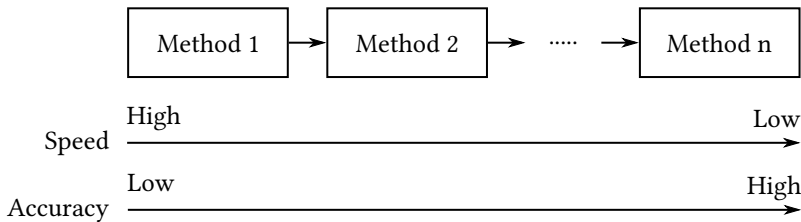


Figure 4.2: A new measurement strategy enabled by AdaScope.

In the following sections, four different methods are investigated for the cascade measurement strategy. Section 4.1 presents the compressive shape from focus method, which is developed for the pre-measurement stage. The target is to perform a fast measurement with a minimum number of camera frames. In Section 4.2, one candidate for the main measurement, namely the iterative array adaptation method, is introduced. Alternatively, the direct area confocal

scanning method is investigated in Section 4.3. For post-measurement refinement, dynamic localized confocal scanning based on Bayesian experimental design is discussed in Section 4.4.

4.1 Compressive Shape from Focus

The fastest measurement mode of AdaScope is achieved when all lateral locations are measured simultaneously, i.e., all pixels of the spatial DMD are turned on. In this case, no lateral scanning is required and the system behaves similarly to a shape from focus setup, where the acquisition of the focal stack can be implemented through scanning of the illumination wavelength.

Estimation accuracy of the conventional shape from focus techniques is strongly coupled with the number of images in the focal stack, thus limiting the measurement speed. In this section, a novel compressive shape from focus method is proposed with an exemplary algorithm based on the modified Laplacian operator (LAPM) and principle component analysis (PCA). Simulation with synthetic focal stacks have demonstrated comparable results to the conventional method. A test with six compressively captured images achieves the same level of performance to that of the conventional method with 100 images. Several other focus measure algorithms are also implemented and tested under the compressive scheme, which demonstrates the wide applicability of the proposed method.

4.1.1 Background

The key concept of recovering depth information from a focal stack is the relationship between focused and defocused images. In a thin lens model, image points that are sharply projected on an image plane fulfill the Gaussian lens equation (Figure 4.3):

$$\frac{1}{d_1} + \frac{1}{d_2} = \frac{1}{f} \quad (4.1)$$

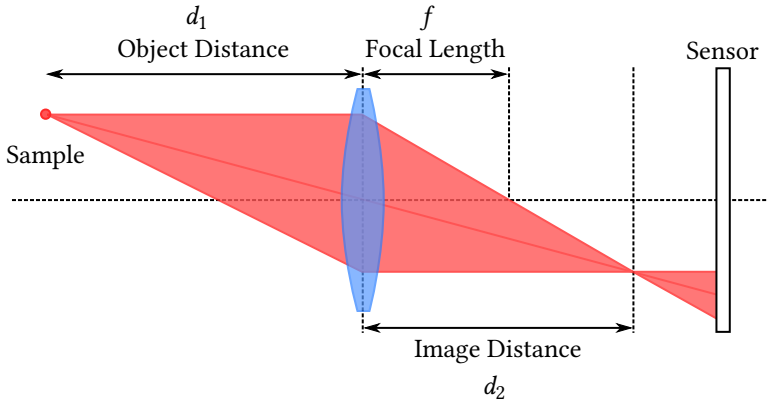


Figure 4.3: Imaging of a thin lens.

where d_1 is the distance of the object point from the lens plane, d_2 denotes the distance of the focused image from the lens plane and f represents the focal length of the optical system. When the detector is moved away from the focus position, the image of the object will be blurred. The degree of blurring depends on how far away the object is from the in-focus position as well as the characteristics of the imaging system.

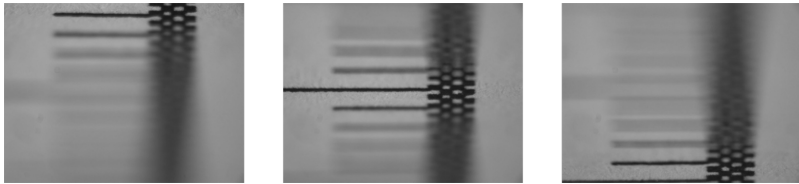


Figure 4.4: Sample images from a focal stack using an imaging system with a small depth of field.

Utilizing the relationship between the blur and the distance to the focus, conventional shape from focus methods are composed of mainly three steps. Firstly, a stack of images are captured while the focus of the imaging system is shifted with respect to the object. This is typically implemented through

either mechanical scanning of the camera/sample, or motorized focus shifting with the lens. Secondly, a focus measure value is calculated for each pixel of every image in the stack to form a 3D focus measure cube, where two dimensions represent the transverse spatial coordinates corresponding to the camera pixels and one dimension denotes the axial shift coordinate. The focus measure value can be calculated with various algorithms to evaluate how well the underlying pixel is in focus. Last but not least, depth information of each pixel is retrieved based on its focus measure curve. Regardless of the focus measure algorithms, the focus measure values for a specific pixel at any axial locations within the measurement range typically forms a Gaussian-shaped signal, which is similar to the axial confocal signal illustrated in Figure 1.2. A naive approach is to take the axial position with maximum focus measure value as the axial position of the object at this lateral location. More sophisticated approaches involve fitting of the focus measure curve as well as optimization techniques, such as total variation regularization, which have been discussed in Section 2.2.

According to estimation theory, the measurement uncertainty is limited by the Cramér-Rao lower bound [van07], which is closely related with the gradient of the signal. For a Gaussian-shaped signal, the maximum gradient is determined by the width of the peak. Since a narrower peak leads to more accurate measurement, shape from focus systems typically aim to achieve a depth of field as small as possible, by using optical systems with larger aperture. Nevertheless, a narrower depth of field requires that the acquisition of the focal stack has to be conducted more densely, which can be very time consuming. To solve this problem, a method for compressive acquisition of the focal stack is developed.

4.1.2 Linear Measurement Model

Various real-world signals can be viewed as an n_1 -dimensional vector $\mathbf{x} \in \mathbb{R}^{n_1}$, including sound, image, etc. In a linear measurement model, each measurement of the target signal is a linear combination of all values in the vector \mathbf{x} .

The complete measurements of the signal can be written as an n_2 -dimensional vector $\mathbf{y} = \mathbf{A}\mathbf{x} \in \mathbb{R}^{n_2}$ with a measurement matrix $\mathbf{A} \in \mathbb{R}^{n_2 \times n_1}$.

The ultimate goal of the linear measurement model, like any other measurement system, is to retrieve the signal \mathbf{x} and the information it is carrying. Formulation of the linear measurement model as a linear system naturally leads to a classical problem of linear algebra: conditions for solving the equation $\mathbf{y} = \mathbf{A}\mathbf{x}$. In this context, this question is equivalent to what kind of measurements are needed in order to recover the signal.

Although prevented by the classical theory of linear algebra, recent developments in compressive sensing have shown that an underdetermined linear system can be uniquely solved provided sufficient prior knowledge [Can06]. In the case of compressive sensing, such prior knowledge refers to the assumption of sparsity. However, this is not the only possible prior knowledge. From a more general perspective, the underdetermined linear system with prior information represents a linear manifold learning problem where the prior information acts as the boundary of the manifold to be learned by its low-dimensional projection. The fundamental philosophy behind solutions of such problems is that the information embedded inside the high dimensional manifold is intrinsically of low dimension. In the case of compressive sensing, the unknown manifold is limited to hyperplanes spanned by a limited number of axes which corresponds to the sparsity assumption.

The significance of the linear measurement model to conventional SFF approaches is that the number of images required in the focal stack can be effectively compressed if each image can act as a linear combination of all originally required images in the focal stack. The focus measure stack can then be retrieved from the focus measure values calculated from the compressed images. It should be noted that this is only possible when the focus measure operator is linear, which is rarely true for modern focus measure operators. Fortunately, most focus measure operators are composed of several sub-operators, and as long as there is at least one linear sub-operator before all nonlinear sub-operators, the reconstruction can be inserted. In other words, the first sub-operator applied on the compressed images must be linear.

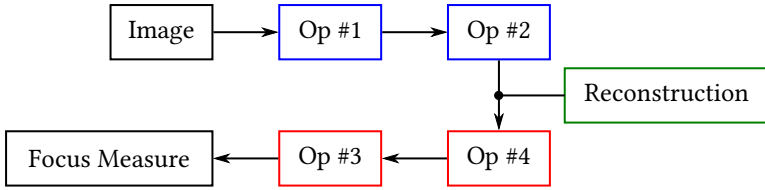


Figure 4.5: The compression and recovery steps must be added before all non-linear operators due to their linear nature. Blue: linear operators. Red: non-linear operators.

The algorithm for the reconstruction of the focus measure stack depends on the prior knowledge, i.e. the focus measure operator. On one hand, when the focus measure curve has a defined peak, recent compressive sensing algorithms can be incorporated for the recovery of the whole curve. On the other hand, if a training process is allowed or assumptions regarding the focus measure curves can be made, conventional methods like PCA can be applied in this scheme to yield the measurement/compressing matrix and the corresponding reconstruction matrix.

4.1.3 Compressive Algorithm

To explain the idea in a concrete and clear manner, an exemplary algorithm is presented in this section. The schematic of the algorithm is illustrated in Figure 4.6.

The measurement matrix forming the compressed images and the reconstruction matrix for decompression are designed by a training process. In this process, conventional SFF procedures are implemented on a sample focal stack so that the focus measure curve for each pixel is calculated. All the focus measure curves are then assembled, with which PCA is conducted. The largest components are combined to construct the measurement matrix for the compressed images and the reconstruction matrix is simply the transpose of the measurement matrix. The focus measure curve can be reconstructed by multiplying the focus measure values of the compressed images with the reconstruction matrix:

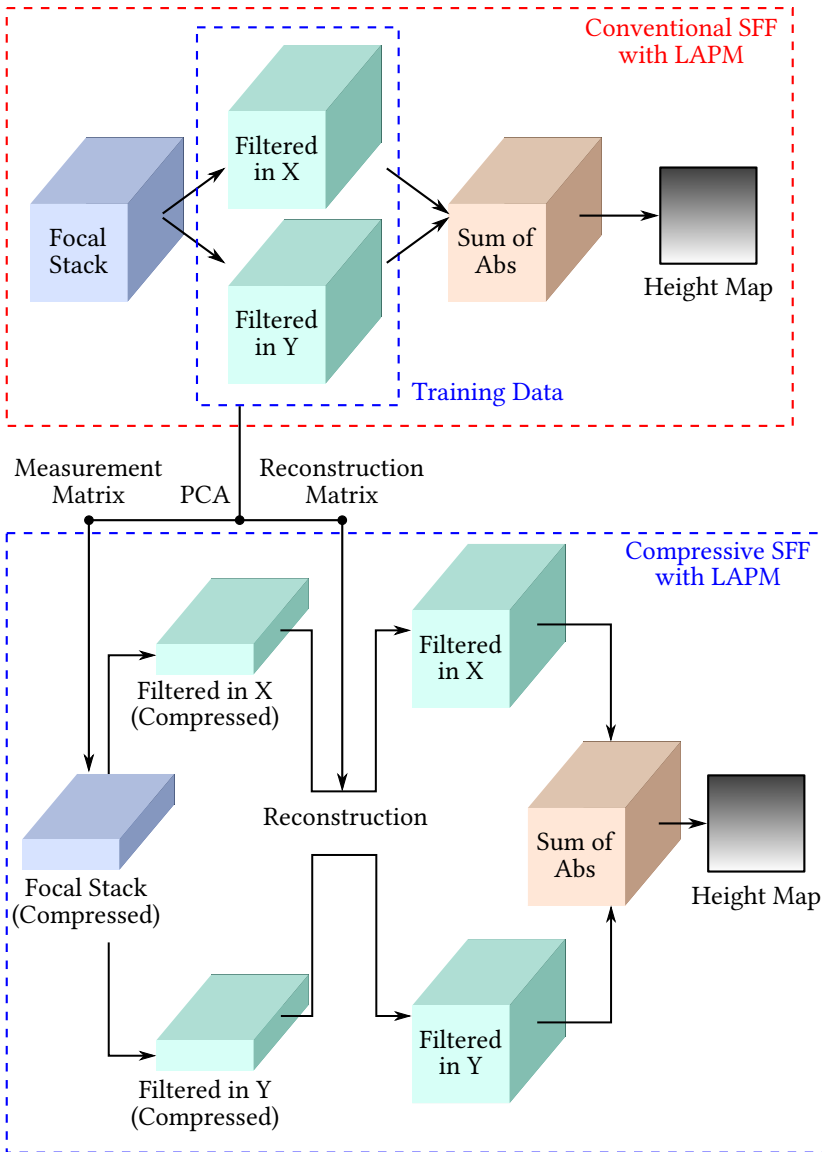


Figure 4.6: Schematic of compressive SFF with the LAPM operator.

$$\mathbf{x}_r = \mathbf{A}^\top \mathbf{y} = \mathbf{A}^\top \mathbf{A} \mathbf{x}, \quad (4.2)$$

where \mathbf{x}_r is the original focus measure curve.

The widely accepted modified Laplacian operator (LAPM) is selected for the calculation of the focus measure [Nay94]. It consists of two sub-operators. Firstly, a one dimensional Laplacian filter is constructed as $\mathbf{f}_{\text{LAP}} = (-1, 2, -1)^\top$ and used to filter the image in both X and Y directions respectively. Secondly, the absolute values of the two filtered images are summed as the final focus measure value. Apparently the 1D filtering operation as a convolution is linear while taking the absolute value is non-linear. Therefore the training and reconstruction step must be inserted before taking the absolute value. From the recovered datacubes of the filtering results in X and Y directions, the final focus measure value can be computed through the sum of the two absolute values. Then for each pixel, the axial focus measure curve is smoothed before the maximum value is located to estimate the axial depth.

4.1.4 Simulation and Discussion

4.1.4.1 Dataset Construction

To demonstrate the applicability of the proposed algorithm, simulation is implemented in Matlab with a series of datasets synthetically generated through programs provided by Pertuz et al. in their survey study [Per13]. The generation of the focal stacks is based on a non-linear shift-variant model of defocus. All estimation results shown in this section are smoothed with a mean filter (window size = 5).

To investigate into the influence of the training dataset on the compressive SFF result, two texture maps and two depthmaps are combined to form four different datasets. Texture #1 is a structured concentric pattern while texture #2 is a random pattern. Depthmap #1 is a linear ramp and depthmap #2 is part of a sphere. The four datasets #1-#4 are formed with combination of textures

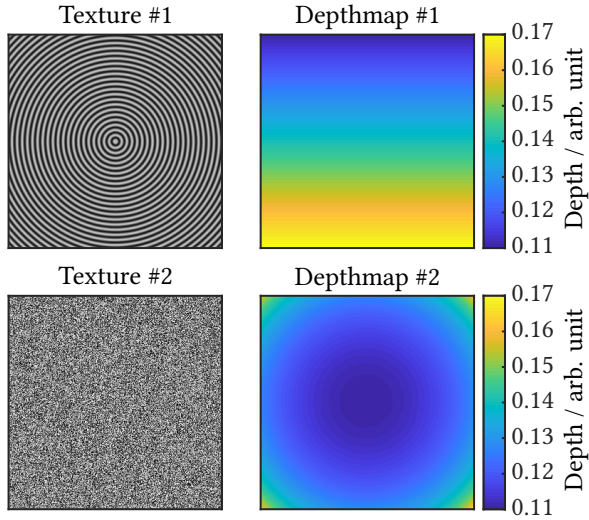


Figure 4.7: Textures and depthmaps used to synthesize focal stacks.

and depthmaps in the following order: #1 and #1, #1 and #2, #2 and #1, #2 and #2.

4.1.4.2 Simulation Result

Table 4.1: RMS error showing influence of training set on testing result.

RMS Error ($\times 10^{-3}$)	Set #1	Set #2	Set #3	Set #4
No Training	1.69	1.74	4.92	3.25
Training Set #1	N/A	0.62	1.19	0.79
Training Set #2	1.11	N/A	2.98	0.72
Training Set #3	4.23	2.40	N/A	0.46
Training Set #4	6.95	2.52	5.58	N/A
Training Set #1-#4	3.65	1.93	1.32	0.54

Results of CSFF are compared with those of conventional SFF using the root-mean-square (RMS) error with respect to the ground-truth depthmaps, which

are listed in Table 4.1. For the conventional method, a focal stack of 61 images are generated in each case and for the CSFF method the 61 images are compressed into 6 images. The row labeled as no training represents the conventional case whereas the other rows are labeled with their corresponding training set, which is used to generate the measurement matrix and the reconstruction matrix. It can be seen from Table 4.1 that the choice of the training set has an influence on the testing result. In general, the compressive results are comparable to the conventional results but require much smaller numbers of compressively captured images. The test result of set #1 with training set #2 and the test result of set #4 with training set #1 are illustrated in Figure 4.8. Several cases in Table 4.1 show that the CSFF method achieves even lower error than the SFF method. This is partly due to the fact that the compression and reconstruction process effectively applies smoothing to the focus measure curve, which is typically quite noisy in the conventional SFF method.

To investigate into the number of images needed for SFF, a series of focal stacks with different numbers of images is synthesized based on texture #1 and depthmap #1 (same combination as dataset #1 used in previous simulations). As expected for the conventional SFF scheme, when the number of images increases, the RMS error decreases, indicating better estimation result. This is due to the fact that the simulated imaging system for image synthesizing has a limited depth of field defined by the blurring kernel. When the step between two adjacent images is too large, the areas with depth in the interval between two focal planes will never get the chance to be imaged sharply and thus cannot be estimated robustly. With the conventional SFF scheme, the minimum number of images needed for robust estimation depends largely on the depth of field of the imaging system, which determines the width of the peak in the focus measure curve when using an operator like LAPM. Generally speaking, when the step size is larger than the width of the focus measure curve, measurement artifacts will start to appear in the estimation result. The dependency of the estimation accuracy on the number of images in the focal stack is illustrated by the blue curve in Figure 4.9. With the current optical configuration, the conventional SFF method reaches the performance limit at approximately 85 images.

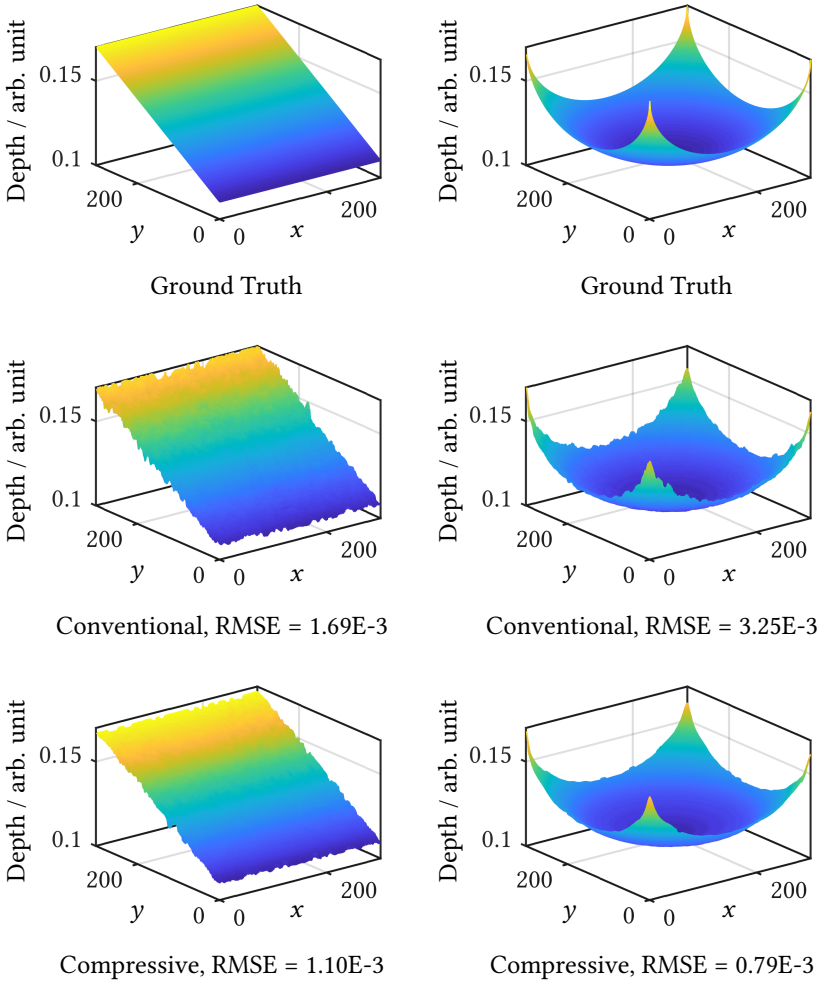


Figure 4.8: Depth estimation results with conventional SFF and compressive SFF. The left column illustrates result of the test set #1 with training set #2 and the right column illustrates result of the test set #4 with training set #1.

On the contrary, in CSFF, regardless of the number of compressive images to be captured, each image acts as a linear combination of all focal planes

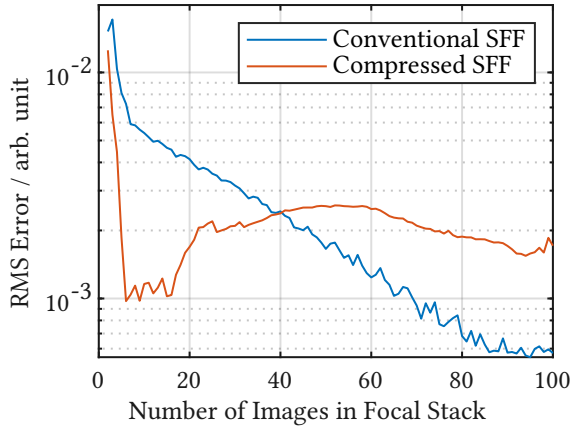


Figure 4.9: Dependency of the estimation accuracy on the number of input images. The error is plotted in logarithmic scale.

within the measurement range, and thus contains information from all focal positions in an encoded manner. A training dataset based on texture #1 and depthmap #2 is synthesized with 100 images. The number of compressive images is solely determined by the number of largest principal components to be selected for the construction of the measurement matrix. As shown in Figure 4.9, CSFF allows much less images to be captured to achieve the same level of estimation accuracy as the conventional method. With a number of 6 images, the estimation performance reaches the limit, which is comparable to the performance of the conventional SFF method with 70 images. As the number of channels increases, the performance drops due to the problem of overfitting. With more than 16 principle components, the measurement matrix is increasingly adapted to the training dataset, resulting in the rise of the RMS error for the testing dataset. It should be noted that results presented above demonstrate the feasibility of the method only on the theoretical level with synthetic datasets. In practice, the performance of both methods could be degraded by various sources of noise.

As the information contained in the largest principal components is related with the rank of the matrix, it is preferred to have a matrix with a relatively

low rank so that more information is contained within a smaller number of principle components. This means that the width of the focus measure curve should be larger, i.e., the imaging system should have a larger depth of field. However, as the width gets larger, the relative magnitude of the focus measure peak gets smaller, effectively reducing the SNR of the measurement. Therefore, a balance must be made between these two factors to achieve the best estimation performance.

4.1.4.3 Applicability

To investigate the applicability of the compressive approach, several other focus measure algorithms are implemented and simulated with the synthetic datasets, including the diagonal Laplacian operator (LAPD) [The09], the Tenegrad algorithm (TENG) [San97] and the steerable filters algorithm (SFIL) [Min09]. Similar to the LAPM, the LAPD also applies Laplacian operators to the captured images, but in two additional diagonal directions. Based on the gradients of the image, the Tenegrad method applies Sobel filters to the image in both directions and then calculates a squared sum. The steerable filters algorithm is a sophisticated modern algorithm that has attracted quite a lot of attention. The focus measure value is calculated using steerable filters in several directions, which are designed in quadrature pairs for better control over phase and orientation. The maximum of the filtered results is taken as the focus measure. Mathematical details regarding these algorithms can be found in the respective literature.

Two groups of tests are conducted using the compressive approach proposed previously with all four focus measure algorithms. For group #1, dataset #1 is tested using the training result from dataset #2. For group #2, dataset #4 is tested using the training result from dataset #1. All algorithms are modified so that the compression and reconstruction processes are inserted before any non-linear operations. As shown in Figure 4.10, all four focus measure algorithms have provided similar results under the compressive scheme. Therefore, the compressive scheme is in general not very sensitive to the selection of the focus measure algorithms as long as the linearity condition mentioned in

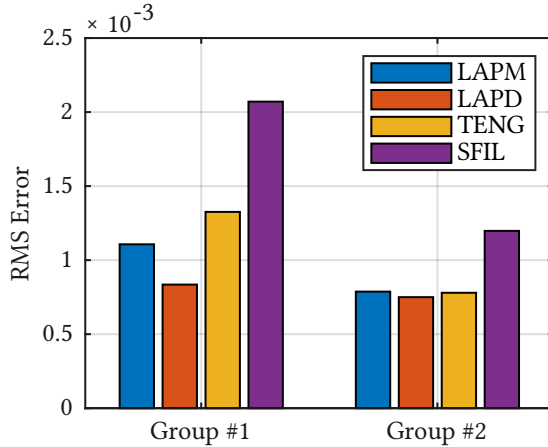


Figure 4.10: Comparison of CSFF using different focus measure algorithms. Group #1: test set #1 with training set #2. Group #2: test set #4 with training set #1.

Section 4.1.2 is satisfied. Nevertheless, for both groups of tests, the steerable filters algorithm has demonstrated noticeably worse results than the other three algorithms. As a more sophisticated algorithm, SFIL should in principle generate better results than the other three algorithms when applied in conventional SFF. However, due to its higher complexity, the added noise introduced by the compression and reconstruction might have more severe influence over the final focus measure calculation, leading to a worse overall result. This implies that the degeneration caused by the compression is possibly more severe for more complicated focus measure algorithms, which has to be taken into consideration when applying compressive shape from focus in practice.

4.1.5 Summary

In this section, a novel method of compressive shape from focus is presented and simulated. Based on the linear measurement model, the CSFF method compressively captures several images, each as a linear combination of all

possible focal planes within the measurement range. It has been shown in the simulation that the estimation error of CSFF is comparable to the conventional method using the same number of images as the number of images in the training set for CSFF. With the synthesized datasets, CSFF with 6 compressive images yields comparable performance to the conventional method with a focal stack of 70 images. Apart from LAPM, several other focus measure algorithms are also tested under the compressive approach, indicating wide applicability of the method.

4.2 Iterative Array Adaptation for 3D Confocal Scanning

As presented by the simulation result in Chapter 4.1, the compressive shape from focus method is able to retrieve the 3D surface profile with a minimum number of images. In practice, the focus measure operators are generally very sensitive to camera noise and the choice of an optimum focus operator depends highly on the surface texture, degrading the robustness of the method in real applications. Nevertheless, measurement accuracy for smooth surfaces is often sufficient to significantly restrict the axial measurement range, so that a more accurate measurement method can be initiated.

In this chapter, an iterative array adaptation method is proposed for confocal 3D scanning. Unlike conventional array scanning methods where a fixed pitch distance is specified, the pitch distance and the axial measurement range are collectively adapted in an iterative manner in order to achieve a higher scanning efficiency.

4.2.1 Motivation and Concept

The idea originates from the observation that the uncertainty of the chromatic confocal measurement is in fact coupled with the lateral density of measurement locations, as shown in Figure 4.11. When little information of the measurement locations is gathered, the crosstalk could potentially be very large

and therefore a larger distance between adjacent points is required. As the measurements at each point become more and more accurate, the possibly generated crosstalk also gets smaller which allows for a denser measurement array.

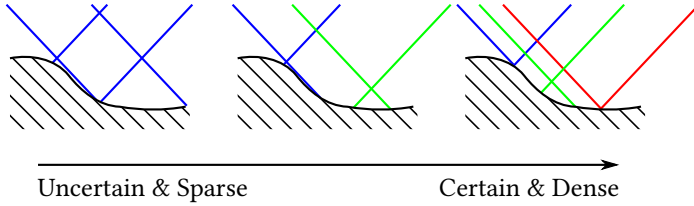


Figure 4.11: Coupling of axial measurement uncertainty and lateral measurement density.

Based on this observation, the measurement is conducted in several iterations. In each iteration, measurements with limited accuracy are made for all positions through array scanning with a fixed pitch distance. Based on the result from one iteration, more refined measurements are made with a denser grid in the next iteration.

4.2.2 Axial Measurement Refinement

In each iteration, a two-channel linear measurement system is applied to a scanning array of measurement locations. The two measurement filter functions are two ramp-shaped functions in opposite directions. To measure the axial location of the corresponding chromatic confocal peak, illuminations with spectra in the shape of the measurement functions are applied and the corresponding images are captured. As Bernstein polynomials of degree 1, these functions have the nice property that the corresponding linear transformation maintains the centroid of the original signal (proof in Appendix A.2). Therefore, the centroid of the chromatic confocal peak can be estimated with very fast computation:

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \mathbf{f}_1^\top \\ \mathbf{f}_2^\top \end{bmatrix} \mathbf{x}, \quad (4.3)$$

$$\text{COG}(\mathbf{x}) = \text{COG}(\mathbf{y}) = \frac{y_2}{y_1 + y_2}, \quad (4.4)$$

where \mathbf{x} represents the original confocal signal, \mathbf{y} denotes the measurement and $\text{COG}(\cdot)$ represents the center of gravity. The illumination spectra are represented by f_1 and f_2 respectively.

There are several reasons for using such a linear measurement system. Firstly, since multiple iterations are performed, each iteration must be very efficient in terms of the number of frames taken. Secondly, the crosstalk at a fixed distance should be proportional to the measurement range. This means that as the location of the object becomes more certain, the crosstalk should become smaller. Lastly, the uncertainty should be inversely proportional to the measurement range. This means that for a smaller measurement range, the sensitivity should be higher.

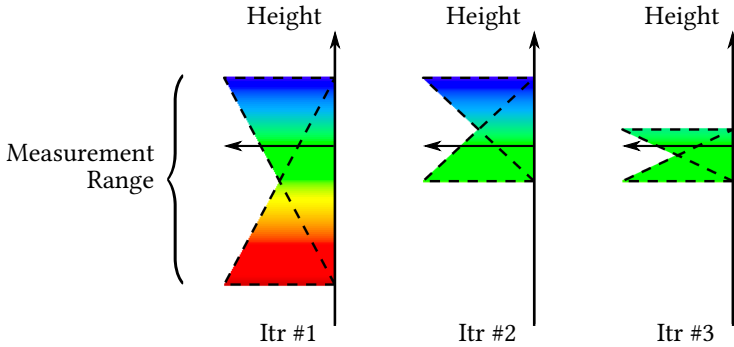


Figure 4.12: Iterative refinement of axial measurement.

All these properties are achieved by iteratively reducing the wavelength range of the illumination according to the previous estimation, such as illustrated by Figure 4.12. The position of the object is represented by the arrow. In the first iteration, the camera takes two frames with the two illumination spectra covering the complete wavelength range. Based on estimation result from the first iteration, which is not extremely accurate, the object is determined to be in the top half of the measurement range. In the second iteration, the

AdaScope makes measurement in the new measurement range with two similar illumination spectra. This appears to be like a binary search, but if the measurement in each iteration is accurate enough, the search process can be much faster. For example, a direct jump from iteration #1 to iteration #3 will also be possible.

Apparently this method is not very sensitive and is not robust against the noise due to the limited number of linear measurement channels, but it should be sufficient to bound the measurement range to a certain level for the next iteration.

4.2.3 Lateral Array Condensation

As mentioned previously, in each iteration, the measurement density is also increased accordingly. As shown by the example in Figure 4.13, in iteration #1 with a pitch of 20 pixels, the point array has to be scanned by 400 times, and in each time, the system makes two measurements using the corresponding illumination spectra. In the next iteration, the density of the array can be increased depending on how much the new measurement range is bounded.

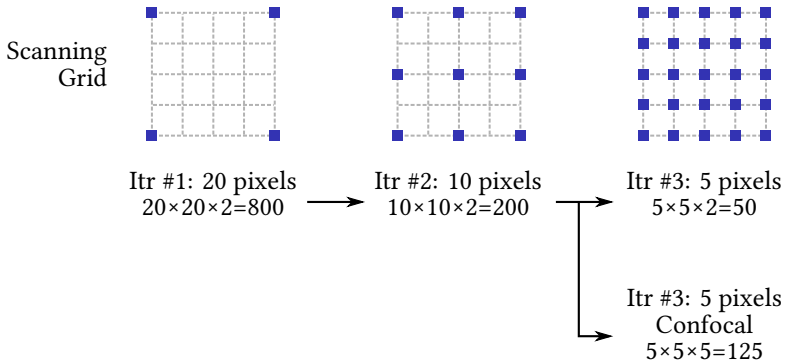


Figure 4.13: Iterative condensation of the lateral array. The DMD pitch distance and the numbers of frames per iteration are labeled respectively.

At a certain iteration, based on the estimation uncertainty from the previous iteration, the measurement process can be switched to a localized chromatic confocal measurement centered around the previous estimation result, in order to get more accurate measurement results.

4.2.4 Triggering Mechanism

As an example, the triggering diagram for the second iteration as well as the corresponding illumination spectra are illustrated in Figure 4.14. The camera serves as the master which triggers the spectral DMD in the programmable light source. This DMD displays several patterns corresponding to several illumination spectra. Each spectral DMD patterns triggers its corresponding spatial DMD pattern in the microscope. Based on the estimation from the first iteration, all points are already bounded to either the top half or the bottom half of the complete measurement range. For each measurement grid, two frames are captured. Within each frame, two spectra are projected to two different spatial patterns. In the second frame, the spatial patterns are repeated but the spectra are different. This process is then repeated n_p^2 times for complete measurement of this iteration, where n_p denotes the pitch distance of the current iteration.

4.2.5 Summary

In this section, an iterative array adaptation method for 3D confocal scanning is proposed. Instead of keeping a constant pitch distance for the array scanning, multiple iterations of array scanning can be performed while the pitch distance and the axial measurement range are modified dynamically. For the axial scanning, a linear measurement method based on the Bernstein polynomials has been developed, where the axial range is halved from iteration to iteration. For the lateral direction, the array density is also increased iteratively. This results in a much more efficient 3D scanning procedure compared to the conventional array scanning method.

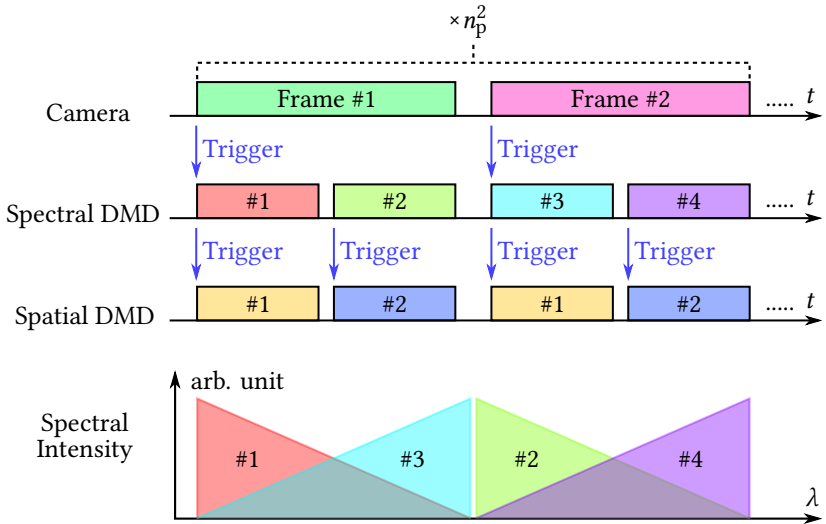


Figure 4.14: Exemplary triggering diagram of iteration #2 and the corresponding illumination spectra.

4.3 Direct Area Confocal Scanning

Although the iterative array adaptation method improves the scanning speed by dynamically changing the axial measurement range and the lateral array pitch, in each iteration, a certain pitch distance between the adjacent measurement locations still has to be guaranteed. Meanwhile, due to the wider bandwidth of the illumination spectra, the crosstalk could still affect the measurement result despite the reduced axial range in each iteration.

In this section, an alternative method is presented based on an entirely different principle, i.e., direct area confocal scanning. This method is both more efficient and more accurate compared to the iterative array adaptation method, and thus proves to be a better choice for the main measurement stage in the cascade strategy.

4.3.1 Theoretical Analysis

Direct area confocal measurement is generally considered impossible due to the fundamental limitations of illumination and imaging. This has been partly circumvented by spectrally encoded slit confocal microscope, where one lateral axis is tackled with a physical slit and the orthogonal lateral axis is covered by lateral dispersion of the slit [Kim06, Kim15]. This section presents an alternative approach for direct area confocal measurement based on a completely different principle, which, as will be demonstrated in the latter sections, provides additional benefits for the complete scanning process in the proposed system.

4.3.1.1 Optical Model

All confocal systems rely on the same principle that unfocused illumination light gets spread to the adjacent area. The reflected light is further filtered by the confocal pinhole, whether it's a physical pinhole, a fiber end or a single pixel, which finally generates the confocal peak in the detected signal.

Figure 4.15 illustrates three different types of microscopes in reflective configuration, where the illumination arm and the detection arm share the same optical system due to the beam splitter. The imaging processes of these microscopes will be investigated in details, with x and y representing the lateral directions and z representing the axial direction.

Figure 4.15 (a) shows a scanning microscope, where a wide-field illumination is projected onto the object and a vanishingly small pinhole is applied before the detector. Commonly referred to as a type-1a microscope, this setup has been shown to be equivalent to a conventional wide field microscope [Wil84]. The intensity response of such a system to a point object can be expressed as

$$I(u, v) = |h(u, v)|^2, \quad (4.5)$$

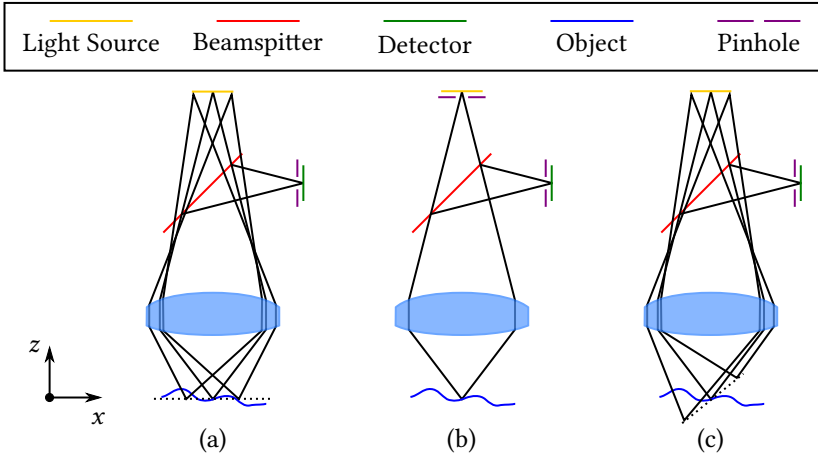


Figure 4.15: Three types of microscopes: (a) Type-1a scanning microscope which is equivalent to the conventional wide-field microscope. (b) Confocal microscope. (c) Type-1a scanning microscope with a tilted illumination field.

where $h(u,v)$ stands for the amplitude point spread function of the optical system with respect to the optical coordinates in the object space:

$$v = kr \sin \alpha = k\sqrt{x^2 + y^2} \sin \alpha, \quad (4.6)$$

$$u \approx 4k\delta z \sin^2 \frac{\alpha}{2}. \quad (4.7)$$

In these equations, k represents the wave number, $\sin \alpha$ is the numerical aperture and δz represents a small axial deviation from the focal plane. For 3D surface profilometry, the more important factor is the integrated axial response of the system, which is denoted by I_{int} . This factor can be treated as the overall intensity in the image of a point object, or approximately considered as the intensity response to a planar diffusing object, which can be written as

$$I_{\text{int}}(u) = 2\pi \int_0^{\infty} I(u, v) v \, dv. \quad (4.8)$$

Although it has been proven that a slit can be used in a confocal system instead of a pinhole with only slightly widened confocal peak [She88], it is apparent that a direct area confocal measurement is not possible. Theoretically, it has been proven with Parseval's theorem that the integrated intensity response in Equation 4.8 does not fall off with respect to u . This can equally be argued with the conservation of energy. When a wide-field illumination is applied to a planar object, all light is reflected and therefore the intensity response remains constant as the object is scanned axially. Consequently, such a system does not possess the capability of depth discerning.

On the contrary, for a confocal system illustrated in Figure 4.15 (b), the intensity response of a point object is given by

$$I(u, v) = |h(u, v)|^4. \quad (4.9)$$

And the integrated intensity response in the focal region can be written as

$$I_{\text{int}}(u) = 2\pi \int_0^{\infty} (C^2(u, v) + S^2(u, v))^2 v \, dv, \quad (4.10)$$

in which $C(u, v)$ and $S(u, v)$ are defined as [Wil84]

$$C(u, v) = \int_0^1 2 \cos\left(\frac{1}{2}ju\rho^2\right)J_0(v\rho)\rho \, d\rho, \quad (4.11)$$

$$S(u, v) = \int_0^1 2 \sin\left(\frac{1}{2}ju\rho^2\right)J_0(v\rho)\rho \, d\rho, \quad (4.12)$$

with J_n as a Bessel function of first kind of order n . The integrated intensity response can be evaluated numerically, which demonstrates that the intensity drops as the object moves away from the focal plane. Such phenomenon serves as the basis for the depth discerning capability of confocal systems. Figure 4.16 illustrates the axial response as well as the integrated intensity for both conventional wide-field microscope and confocal microscope.

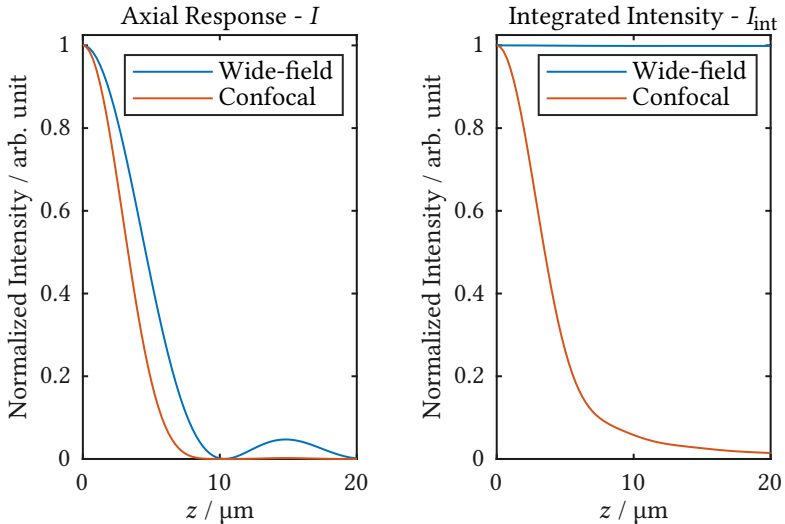


Figure 4.16: Axial response I and integrated intensity I_{int} for wide-field microscope and confocal microscope.

Another system is presented in Figure 4.15 (c), which is similar to a type-1a scanning microscope illustrated in Figure 4.15 (a). The only difference is that the focal plane of the optical system is rotated with respect to the y -axis. One specific implementation of the system is discussed in the following sections and currently the optical system is treated as a black box which is capable of generating such tilted focal field. Since this configuration breaks the radial symmetry of the system, Cartesian coordinates are used instead of the optical coordinates. For incoherent imaging, the intensity response of the system to a point object at (x,y,z) can be expressed as

$$I(x, y, z) = \iiint_{-\infty}^{\infty} H(x - x', y - y', z - z') H(x, y, z - z') F(x', y', z') dx' dy' dz', \quad (4.13)$$

where $H(x,y,z)$ is the intensity point spread function of the optical system and $F(x,y,z)$ is a 3D mask function which defines the illumination distribution.

The integration shown in Equation 4.13 has a form similar to a convolution, where the first H is shifted three-dimensionally and multiplied by the mask function to account for the tilted area illumination field. The second H represents the imaging intensity point spread function by the point object. In the particular implementation with the AdaScope system, chromatic encoding is applied to achieve the tilted illumination field, where the focal length of the optical system varies according to the wavelength. To account for this effect, the axial coordinate z in the second H is also shifted so that the focal length matches that of the illumination wavelength.

With the on axis focal point position defined as $(0,0,0)$ and the angle between the illumination plane and x -axis denoted as θ , the mask function can be expressed as $F(x,y,z,\theta) = \delta(z - x \tan \theta)$. And therefore, Equation 4.13 can be simplified as

$$I(x, y, z, \theta) = \iint_{-\infty}^{\infty} H(x-x', y-y', z-x' \tan \theta) H(x, y, z-x' \tan \theta) dx' dy'. \quad (4.14)$$

The corresponding integrated intensity response can be expressed as

$$I_{\text{int}}(z, \theta) = \iint_{-\infty}^{\infty} I(x, y, z, \theta) dx dy. \quad (4.15)$$

This expression does not have an analytic solution and thus must be evaluated numerically.

4.3.1.2 Simulation result

In this section, Equation 4.14 and Equation 4.15 are evaluated through numerical simulation to investigate its corresponding depth discerning capability for 3D measurement. The intensity point spread function $H(x, y, z)$ is simulated based on a fast 3D PSF model [Li17] for a volume of $80 \mu\text{m} \times 80 \mu\text{m} \times 150 \mu\text{m}$. The numerical integration is implemented afterwards. A numerical aperture of 0.33 is utilized, which corresponds to the experimental setup and a wavelength of 580 nm is specified.

To understand the simulation results, a simpler case with a tilted focal field under slit illumination is firstly investigated. In this case, a slit along the x -axis is applied to the light source generating a tilted line of focused illumination, which forms an angle of θ with respect to the x -axis. The integrated intensity response of such a configuration is simulated and illustrated in Figure 4.17. For each angle θ , the intensity response is normalized so that the maximum intensity equals to one. The bright green curve on the bottom represents the integrated intensity response of a conventional single point confocal microscope. As can be seen from the simulation result, for most of the angles, the

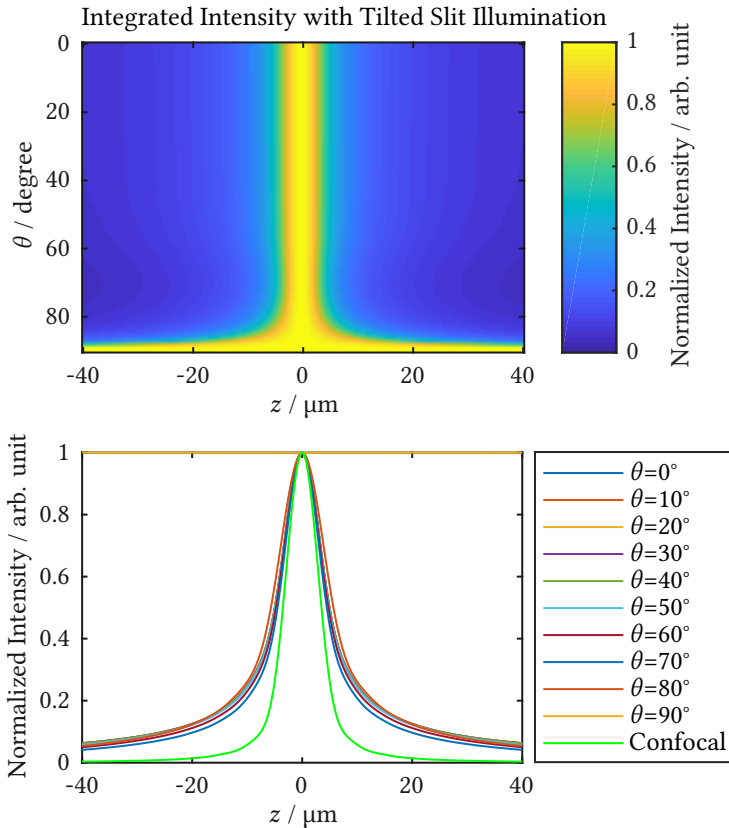


Figure 4.17: Simulation result for tilted slit illumination. Top: axial integrated intensity response as an intensity map. Bottom: axial integrated intensity responses for a series of tilting angles in comparison to the confocal signal.

system remains capable of depth discerning, although the width of the intensity is slightly broadened. It is worth noting that at 0° , the system is exactly a conventional slit scanning confocal system. Whereas at 90° , the system becomes a physically prohibited confocal system with a focal point infinitely elongated along the optical axis. Alternatively, it can be seen as a single point

chromatic confocal system, with a detector not capable of wavelength discerning. Therefore, as the angle approaches 90° , the intensity response becomes more flat and the system gradually loses its depth discerning capability.

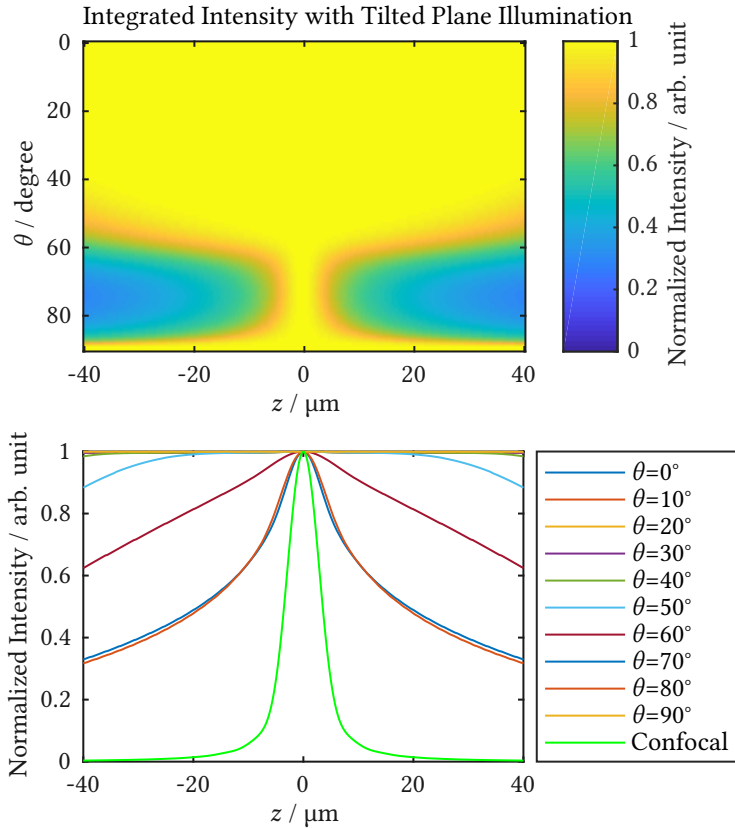


Figure 4.18: Simulation result for tilted plane illumination. Top: axial integrated intensity response as an intensity map. Bottom: axial integrated intensity responses for a series of tilting angles in comparison to the confocal signal.

What's more interesting is to look at the case where a complete planar light source is used to generate a tilted planar illumination field, which is presented in the previous section. As shown by the results illustrated in Figure 4.18, at

0° , the system represents a type 1a scanning microscope, which is equivalent to a wide-field microscope. At 90° , the system can be treated as an array of chromatic confocal point sensors with broadband detectors or simply as a chromatic confocal slit scanning system with a broadband detector. It is apparent that both the case of 0° and the case of 90° represent a system which is not capable of 3D measurement. However, as the angle varies between 0° and 90° , an interesting region arises around roughly 75° , where an intensity peak is clearly visible, indicating the capability of depth discerning. The FWHM of the conventional integrated confocal intensity peak is roughly $7\ \mu\text{m}$ and the FWHM of the peak at 75° is roughly $32\ \mu\text{m}$. Although the intensity peak is several times wider than a conventional confocal peak, this sacrifice leads to an imaging system capable of true area confocal scanning.

Figure 4.19 illustrates the intensity response of the proposed system when a point object is scanned laterally. This is simulated by calculating $I(x,y,z,\theta)$ in Equation 4.14 through numerical integration. In the y -direction, since the plane of illumination is tilted with respect to the y -axis, the response is very similar to that of a slit scanning confocal microscope in the direction parallel to the slit. In the x -direction, the width of the intensity response is less affected by the area illumination but is more sensitive to the change of the tilting angle. In both cases, the FWHM of the signals are slightly wider than a conventional confocal signal, indicating a good lateral resolution very similar to a conventional confocal system.

Based on the simulation results illustrated in Figure 4.18, the axial FWHM for the integrated intensity response can be calculated. Figure 4.20 demonstrates the change of axial FWHM with respect to the tilting angle for three different NAs. Similar to a conventional confocal scanning system, the effective axial FWHM is reduced as the NA is increased. Additionally, a larger NA allows for a larger operational tilting angle range, which moves toward a lower tilting angle. Consequently, the optimum tilting angle for the minimum FWHM is also reduced as the NA increases. For an NA of 0.33, which will be utilized in the experimental setup, the optimum tilting angle is approximately 75° .

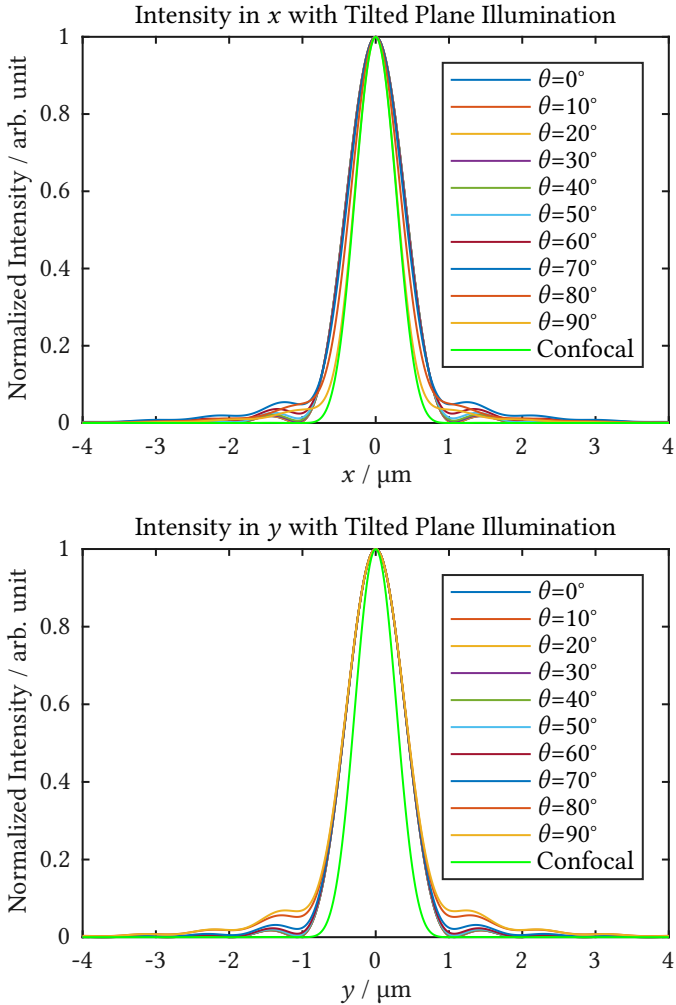


Figure 4.19: Intensity response of a point object with tilted plane illumination. In both x - and y -directions, the intensity response is only slightly wider than a conventional confocal signal.

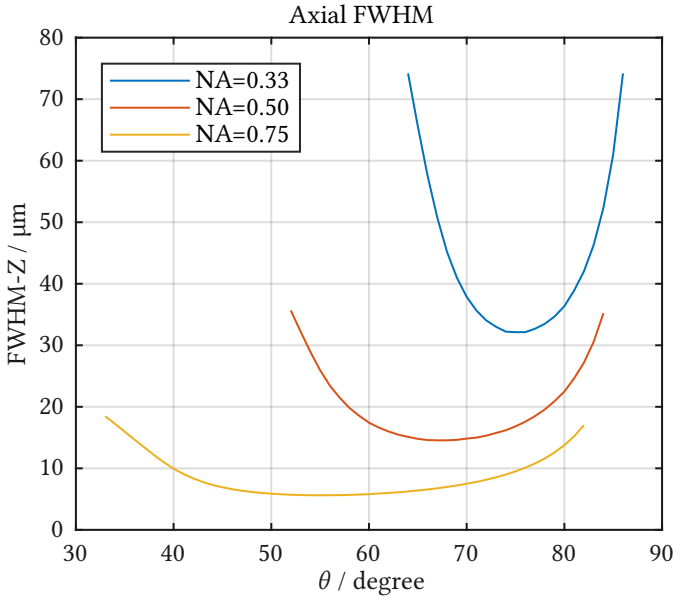


Figure 4.20: Axial FWHM of the integrated intensity with respect to NA and tilting angle θ . For each NA, an optimum angle can be located where the axial FWHM is minimum.

4.3.2 Scanning Mechanism

To implement the method presented in Section 4.2.1 based on the proposed setup, several important aspects are studied and described in detail in this section.

4.3.2.1 Illumination Generation

As demonstrated by the simulation result in Section 4.2.1, the depth discerning capability of the tilted area confocal scanning method is only maintained in a small range of angles around 75°. Although there are possible solutions based on special optical design (e.g. with a Scheimpflug configuration), the imaging quality could be adversely affected. Therefore, in the AdaScope setup,

the tilted focal field is implemented through multiplexed chromatic encoding. Since the illumination spectrum can be tuned through the first DMD in the programmable light source and the lateral locations can also be arbitrarily addressed through the second DMD, the combination of the two DMDs allows the generation of focused illumination field anywhere within the three dimensional measurement volume. Through time multiplexing, within the exposure time of each camera frame, several pairs of patterns are displayed by the two DMDs, forming a series of localized illumination distributions in the target space. By controlling the two DMDs simultaneously, it is possible to generate any 3D focal field, including the required tilted planar illumination field. In this setup, the programmable light source generates a series of spectral Gaussian peaks of equal intensity with a FWHM of 1 nm. Each spectral peak corresponds to one column of lateral DMD pixels. The tilting angle of the illumination plane can be calculated through the following expression:

$$\theta = \arctan\left(\frac{\Delta\lambda \times a_{z,\lambda}}{d_p \times M_1}\right), \quad (4.16)$$

where $\Delta\lambda$ represents the wavelength step between the adjacent spectral peaks and M_1 denotes the magnification of the illumination arm, which equals to 0.37. The chromatic focal shift $a_{z,\lambda}$ can be expressed as $a_{z,\lambda} = \delta z / \delta\lambda$. And d_p denotes the physical pitch of a single lateral DMD pixel, which is $7.56 \mu\text{m}$ in the proposed setup. For example, at a wavelength of 530 nm, the chromatic focal shift is approximately $29.13 \mu\text{m nm}^{-1}$. To generate a tilting angle of 75° , a wavelength step of 0.36 nm is required between adjacent columns of pixels. Due to the nonlinear chromatic focal shift (Figure 3.24), to maintain a linear axial spacing, the wavelength step between adjacent DMD pixel columns must vary according to the corresponding wavelength. Such implementation of tilted planar illumination field differs from the idealized theoretical model from Section 4.2.1 mainly in two aspects. Firstly, while in the model the illumination plane is continuous and extends in all dimensions to infinity, in practice it is apparent that such illumination is impossible and is approximated by a finite illumination plane composed of discrete illuminated locations. Secondly, the numerical aperture of the different wavelengths varies

slightly instead of remaining a constant value, which is assumed in the theoretical model. Despite these discrepancies, the theoretical model is considered a valid approximation to the practical setup.

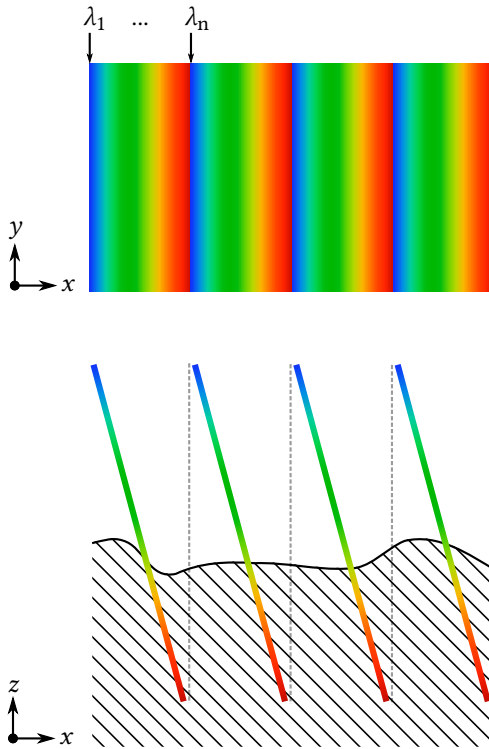


Figure 4.21: Periodic planar illumination with the adaptive microscope. The tilting of the illumination field is achieved through the lateral change of the illumination wavelength and the axial chromatic focal shift of the optical system.

Due to the large tilting angle, even for the full axial range of 4.67 mm, the effective lateral coverage of illumination in x direction is only 1.25 mm. This seems to be a major drawback of the tilted area scanning method. Nevertheless, thanks to the adaptability of the proposed system, multiple periods of illumination planes can be easily configured (Figure 4.21). Although the

boundary area of two adjacent illumination periods may be susceptible to additional crosstalk, as will be shown by the experimental result, the adverse influence is well within acceptable level.

4.3.2.2 Scanning Direction

The confocal scanning is achieved through manipulation of the 3D illumination field based on the control of the two DMDs, while the camera records one image for each illumination field. As the complete illumination is shifted axially, part of the illumination field which is out of the measurement volume will be wrapped in from the opposite side. For example, Figure 4.22 demonstrates the course of a scanning process. Since the tilting angle of the illumination field is much larger than 45° , it is more intuitive to consider that the illumination field is being scanned laterally along the x -axis. In fact, due to the two dimensional relative movement and the wrapping of the illumination field, scanning in the x -direction is exactly equivalent to scanning in the z -direction with a different scanning speed. One additional benefit brought by such equivalency is that the direct area confocal scanning method is potentially applicable on an assembly line, where one period of tilted illumination pattern can be fixed while the object is scanned along one lateral direction by the transporting system.

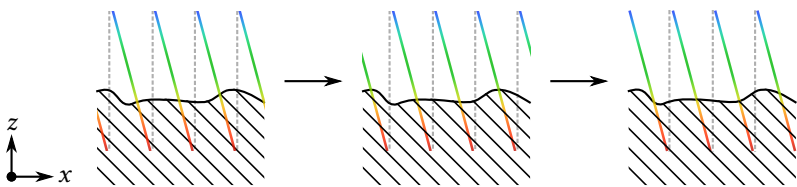


Figure 4.22: The illumination field is scanned by shifting the illumination wavelengths laterally.

The total number of images can be adaptively tuned according to the axial measurement range and the discretization of the illumination field, which depends on the required accuracy of the measurement. For example, if 50 illumination spectra are utilized, which corresponds to 50 columns of pixels on

the spatial DMD, a total of 50 images shall be required for a complete three dimensional scan of the surface. By enlarging the FWHM of the spectral peak, the axial intensity response is effectively widened due to an overlapping of multiple shifted illumination planes. This reduces the axial resolution of the system but allows faster scanning, e.g., shifting of the 3D illumination field by multiple pixel columns instead of one column. This property further enhances the adaptability of the measurement system.

4.3.2.3 Synchronization Mechanism

As discussed in Section 3.2.5, the sCMOS camera (Andor Zyla 5.5) is the slowest component in the system and is thus applied as the master in the synchronization in order to fully utilize its potential. The synchronization mechanism is adapted from the mode #1 illustrated from Figure 3.27, where the camera triggers a series of spectral patterns, each of which then triggers its corresponding spatial pattern.

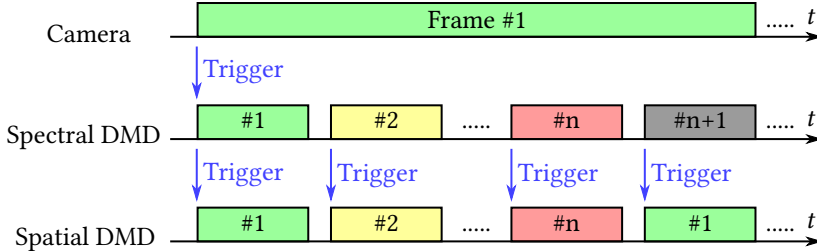


Figure 4.23: Triggering diagram. The camera triggers the spectral DMD to start a series of illumination spectra. Each spectrum triggers a corresponding pattern on the spatial DMD.

To avoid the transmission of new patterns to the DMD between camera frames, a special mechanism is invented by adding an additional black pattern for the spectral DMD at the end. As demonstrated by the triggering diagram in Figure 4.23, at the beginning of each exposure, the camera sends a trigger signal to the spectral DMD, which starts a continuous series of $n + 1$ patterns. Each of the first n spectral DMD patterns corresponds to the

illumination of a spectral peak and triggers its respective spatial pattern. The last spectral DMD pattern is completely black, which serves to send one additional trigger to the spatial DMD. Since the spatial DMD is only loaded with n patterns, this additional trigger from the spectral DMD effectively shifts and wraps one spatial pattern till the end of the next exposure, i.e., the order of the spatial pixel columns is shifted by one pixel. The combined effects result in a shifted 3D periodic illumination field from frame to frame, as illustrated in Figure 4.22. In this way, all patterns can be transferred to the DMDs off-line, and during the measurement, the camera can be operated in a continuous burst mode with its maximum speed.

4.3.3 Summary

Despite the fact that a wide-field microscope lacks the capability of depth discerning, it has been demonstrated in this section that direct area confocal scanning is indeed possible as long as the focal field is tilted to an angle specific to the NA of the system. The simulation results show that the axial confocal response can be largely preserved, yielding an axial FWHM of $32\ \mu\text{m}$ for an NA of 0.33 at the optimum tilting angle. The lateral response is slightly wider than that of the confocal case, demonstrating a good lateral resolution. Compared to the conventional array scanning, the measurement speed is greatly improved as all lateral locations are scanned simultaneously.

4.4 RNN-accelerated Experimental Design

Direct area scanning based on the tilted focal field leads to a significant improvement of the scanning speed compared to conventional array scanning. Nevertheless, measurement uncertainty becomes slightly worse due to the widened signal peak, as shown in Section 4.2.1. To reach the same level or even surpass the accuracy of conventional confocal array scanning, a final stage of refined measurement is required, which is implemented through a localized axial confocal scan based on a fixed lateral array with a significantly smaller pitch distance. Based on the result of the previous measurement stage,

the scan can be initialized and performed directly in the vicinity of the peak position.

Although the localized axial scan can be achieved in a uniform sampling manner, this section introduces a more efficient scanning method based on Bayesian experimental design, which can be further accelerated through an approximation based on a recurrent neural network.

4.4.1 Chromatic Confocal Signal

The target of chromatic confocal measurement is to retrieve the depth of the measurement position via the location of the Gaussian-like signal peak. From the point of view of parameter estimation, the canonical approach is to build a measurement model and apply Bayesian inference on the parameters of interest. The measurement model is composed of two parts, i.e., the signal model and the noise model. The signal model describes the relationship between an ideal signal, or expectation of the signal, and the parameters to be estimated. And the noise model represents the amount of noise added to the ideal model.

In the case of the confocal measurement, the axial intensity response can be derived from Equation 4.9, which has an analytical form:

$$I(u, 0) = \left(\frac{\sin u/4}{u/4} \right)^4. \quad (4.17)$$

As the object is positioned axially in a chromatic confocal system, the detected signal can be approximated by a Gaussian function, which can be expressed by the following equation:

$$\hat{g} = \theta_2 e^{-\frac{(\lambda - \theta_1)^2}{2\sigma^2}}, \quad (4.18)$$

where θ_2 represents the amplitude of the signal and θ_1 represents the location of the signal. The parameter θ_2 is mainly determined by the reflectance

of the object and θ_1 reflects the axial position of the object. The width of the Gaussian-shaped chromatic confocal signal is related with σ and is determined by the properties of the optical system such as the numerical aperture. Assuming normally distributed noise, the complete model is expressed as a normal distribution over the combination of the signal and the noise:

$$g \sim \mathcal{N}(\hat{g}, \sigma_n^2), \quad (4.19)$$

where σ_n^2 describes the variance of the noise and is mainly determined by the camera.

Based on Bayes' theorem, the parameter estimation task is relatively straightforward by calculating the posterior probability distribution of the parameters based on the measurement model. In this case, the parameters of interest are $\boldsymbol{\theta} = (\theta_1, \theta_2)$, where θ_1 contains the depth information and θ_2 contains information of the object texture. The posterior is proportional to the product of the prior and the likelihood. Without any prior knowledge, the prior distribution is considered to be flat across the valid support so that all parameter values are equally possible when no measurements are made. The likelihood comes directly from the measurement model, as shown in Equation 4.19. Therefore, the posterior distribution can be calculated up to a certain scale factor:

$$p(\boldsymbol{\theta}|g) = \frac{p(\boldsymbol{\theta})p(g|\boldsymbol{\theta})}{p(g)} \quad (4.20)$$

$$\propto p(\boldsymbol{\theta})p(g|\boldsymbol{\theta}). \quad (4.21)$$

In practice, the calculation of the posterior distribution with high resolution is often computationally prohibitive, and therefore sampling techniques such as Markov-Chain Monte Carlo (MCMC) method are frequently adopted. For parameter estimation of the chromatic confocal signal, an ensemble sampler

which is affine-invariant [Goo10] is utilized for drawing the posterior samples. Once samples are drawn from the posterior distribution, the estimation becomes trivial by calculating the average of all samples.

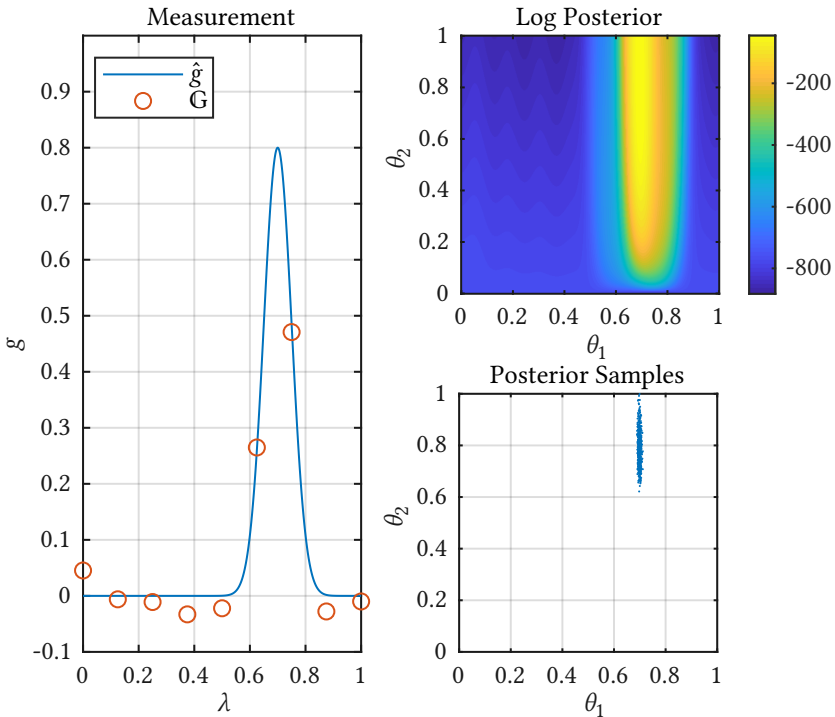


Figure 4.24: Posterior sampling after measurements are made. Both the signal amplitudes (g and θ_2) and the wavelengths (λ and θ_1) are shown in a normalized range from 0 to 1.

Figure 4.24 demonstrates the procedure of posterior sampling for the chromatic confocal measurement through simulation. In the left figure, the expectation of the signal is denoted by \hat{g} and the simulated measurements with normally distributed noise are contained in G . The top right figure illustrates the posterior probability distribution of the parameters to be estimated and the bottom right figure shows samples drawn from such a distribution.

The Bayesian framework has two major advantages for parameter estimation. Firstly, the uncertainty of the estimation can be easily derived by calculating the variance of the samples. Secondly, the posterior distribution of the parameter allows for the selection of the optimal measurement location in the next measurement through Bayesian experimental design, as will be discussed in the following.

4.4.2 Bayesian Experimental Design

Bayesian Experimental Design (BED) is the subject of making decisions under uncertainties with limited resource. In the case of measuring a chromatic confocal signal, conventional systems utilize a spectrometer which disperse various wavelengths onto hundreds of pixels. A major drawback for such approach is that the transfer of the intensity data can be quite slow. Additionally, in the case of an array chromatic confocal system, the application of multiple spectrometers is often prohibitive, due to either cost or mechanical constrains. Therefore, wavelength scanning of the light source is used instead to acquire the chromatic confocal signal. Nevertheless, such process can be time-intensive depending on the scanning speed of the light source.

Instead of an equidistant measuring scheme, Bayesian experimental design allows for an adaptive measuring scheme, where the location for a new measurement is determined by measurements already conducted. For example, when the intensities of several wavelengths have already been measured, the question that BED attempts to answer is which wavelength should be measured next so that the estimation could be made most efficiently. Such approach fits naturally to the post-measurement refinement of the AdaScope, as individual localized axial positions can be scanned dynamically.

The profit generated by a new measurement at a certain wavelength is described by a utility function over the design space. There are various different utility functions which focuses on different aspects of the design. Chaloner and Verdinelli [Cha95] have presented an overview of Bayesian optimal design and discussed appropriate choices for the utility function. For parameter estimation, a common choice is the expected Shannon information gain. The

additional information gained through the new measurement is represented by the Kullback-Leibler (KL) divergence between the current posterior distribution and the updated posterior distribution after the new measurement. The utility function is expressed as the expectation of the KL divergence under the posterior predictive distribution:

$$\begin{aligned}
 U(\xi) &= E_{g|\xi} [D_{\text{KL}}(p(\theta|\mathbf{G}, g, \xi) \| p(\theta|\mathbf{G}))] \\
 &= \iint p(\theta|\mathbf{G}, g, \xi) \log \frac{p(\theta|\mathbf{G}, g, \xi)}{p(\theta|\mathbf{G})} d\theta p(g|\mathbf{G}, \xi) dg \\
 &= \iint p(\theta|\mathbf{G}) p(g|\theta, \xi) \\
 &\quad \left[\log p(g|\theta, \mathbf{G}, \xi) - \log \left[\int p(\theta|\mathbf{G}, \xi) p(g|\theta, \mathbf{G}, \xi) d\theta \right] \right] d\theta dg,
 \end{aligned} \tag{4.22}$$

where ξ represents the possible designs, i.e., the next wavelength to be measured.

Calculation of the double integral for this utility function cannot be conducted analytically and therefore is solved by a nested Monte Carlo (MC) approximation using posterior samples drawn for parameter estimation [Rya03].

$$U(\xi) \approx \hat{U}_{N,M} \equiv \frac{1}{N} \sum_{i=1}^N \left[\log p(g^i|\theta^i, \xi) - \log \left[\frac{1}{M} \sum_{j=1}^M p(g^i|\theta^{ij}, \xi) \right] \right], \tag{4.23}$$

where $\{\theta^i\} \cup \{\theta^{ij}\}$ are drawn from $p(\theta|\mathbf{G})$ and $\{g^i\}$ are drawn from $p(g|\theta^i, \xi)$. The numbers of samples to be drawn are controlled by N and M .

Finally, the task is to find the ξ^* which maximizes the utility function above.

$$\xi^* = \arg \max_{\xi \in [0,1]} U(\xi) \tag{4.24}$$

Although there are stochastic optimization techniques for such a problem, the utility function is typically calculated for a grid of discrete design points and the design point with the highest utility is selected for the next measurement. Notice that this approach is based on the so-called myopic design. It means that only one further step is considered based on the current situation. This does not guarantee true optimal design for an experiment with multiple measurements, but in general works very well as a greedy method.

As an example, Figure 4.25 demonstrates the adaptive measurement of a chromatic confocal signal. The first column shows the signal to be measured and the corresponding measurement in each step. In these graphs, \hat{g} denotes the signal to be measured, g' represents the new measurement in each step and G contains all measurements conducted previously. The second column shows the utility function over the design space in each step. In this example, measurement starts by recording intensity of the wavelength in the middle. Based on the measurement result, parameter estimation is conducted and the utility function over all wavelengths is calculated. In the next step, intensity is measured at the wavelength which has the largest utility value. These two steps can be repeated multiple times until the utility function become close to zero for the whole design space, indicating that new measurements no longer bring any additional information. The wavelength is normalized to a range from zero to one as the calculations are all based on simulations.

Figure 4.26 shows the comparison between the uniform measurement scheme and the adaptive measurement scheme based on BED. As seen from the posterior samples, with the same number of measurement steps, the adaptive approach typically generates much more concentrated samples, indicating less uncertainty for the parameter estimation. The reason is that the adaptive approach tend to make new measurements at locations where more information is expected to be gained regarding the parameters.

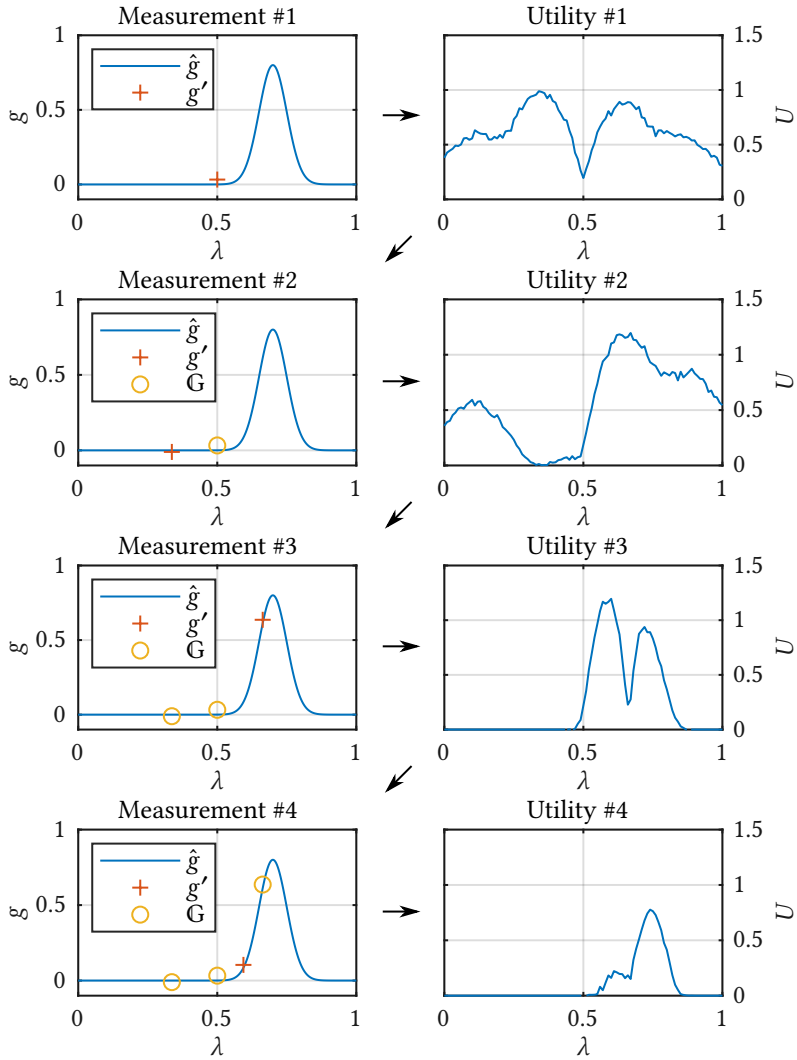


Figure 4.25: Adaptive measurement of a chromatic confocal signal. Left column: each measurement step. Right column: utility function after each measurement step.

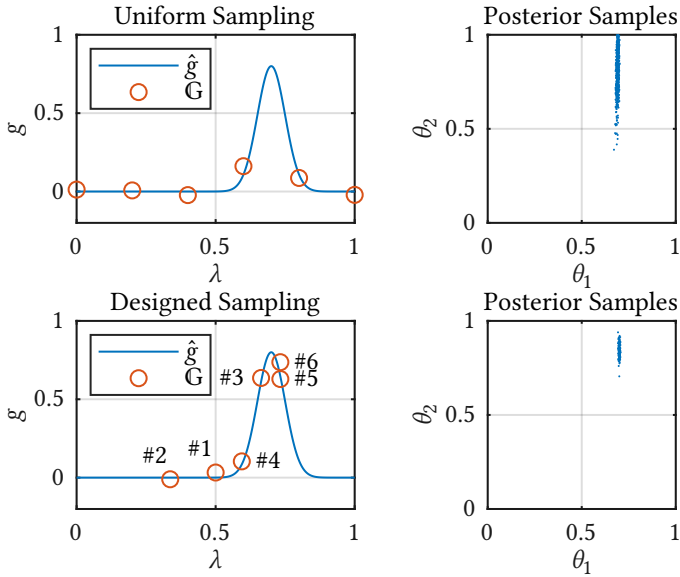


Figure 4.26: Comparison between uniform measurement and adaptive measurement. First row: uniform measurement and its corresponding posterior estimation. Second row: adaptive measurement and its corresponding posterior estimation.

4.4.3 RNN-based Acceleration

As discussed in Section 4.4.2, the utility function in Bayesian experimental design can be approximated by a nested Monte Carlo method shown in Equation 4.23. One major disadvantage of this approach is its slow speed. The nested MC approximation of the utility function shown above is only asymptotically unbiased as an estimator of the utility function. The bias and the variance of the estimator depends on the number of posterior samples. As shown in previous study [Rya03], the variance can be represented as $A_1(\xi)/N + A_2(\xi)/(NM)$ and the bias can be represented to the leading order by $A_3(\xi)/M$, where A_i are terms depending on the sampling distribution and M and N control the numbers of samples to be drawn in the nested MC procedure as defined in Equation 4.23. The number of samples needed for experimental design is naturally much larger than that for pure inference. To make things

even worse, the inner loop of this nested MC has to be performed for each design candidate individually. Due to these reasons, even with faster computers nowadays, full Bayesian experimental design is only implemented in limited fields, such as pharmaceutical studies and astronomy. What's common about these fields is that although the model behind is often very complex, the time interval between two experiments are also very long, thus allowing a good design to be found in a Bayesian way.

Dynamic localized measurement of the chromatic confocal signal is exactly the opposite. Real-time decisions have to be made based on a relatively simple model. If the design speed is not fast enough, it would be more efficient to simply scan the whole wavelength range like a spectrometer. To accelerate the Bayesian experimental design process, a specific type of neural network, i.e., the recurrent neural network, can be trained as an approximation.

The inspiration for using this model originates from a recent topic in computer vision community, namely the Visual Attention Model [Ba14]. For pattern recognition task, the researchers try to mimic the human vision system using a recurrent network. Instead of performing classification on the complete image, a small image patch is processed by the RNN and the output is both the classification result and where to look next. The training is implemented with reinforcement learning. It seems quite obvious that the visual attention model and Bayesian experimental design share an incredible amount of similarities as both attempt to gain more information through a series of adaptive measurement/observation.

For a conventional feed-forward neural network with a single hidden layer, the propagation of data can be expressed as:

$$\begin{aligned} \mathbf{s} &= f_s(\mathbf{W}_s \mathbf{x} + \mathbf{b}_s) \\ \mathbf{o} &= f_o(\mathbf{W}_o \mathbf{s} + \mathbf{b}_o) \end{aligned} \tag{4.25}$$

where \mathbf{x} denotes the input signal, \mathbf{s} and \mathbf{o} represent the activation of the hidden layer and the output layer respectively. \mathbf{W}_s and \mathbf{W}_o are matrices containing

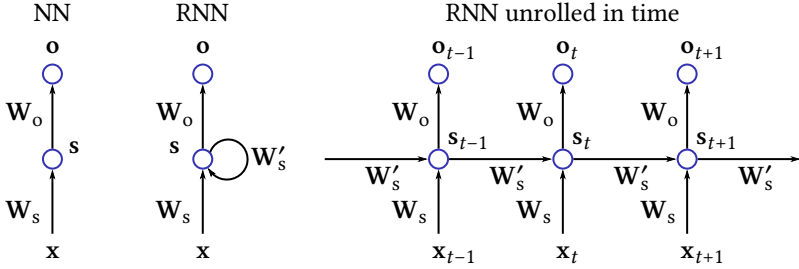


Figure 4.27: Graph representations of the feed-forward neural network and the recurrent neural network.

weights describing the connections between the layers. The non-linear activation functions are represented by f_s and f_o , which can have various forms. And \mathbf{b}_s and \mathbf{b}_o stand for the biases. More layers can be added to form more complex networks.

A recurrent neural network is capable of “memorizing” the previous input data due to the introduction of a feedback loop in the hidden layer. Although more sophisticated variations have been developed, the simplest form of an RNN can be expressed as:

$$\begin{aligned} \mathbf{s}_t &= f_s(\mathbf{W}_s \mathbf{x}_t + \mathbf{W}'_s \mathbf{s}_{t-1} + \mathbf{b}_s) \\ \mathbf{o}_t &= f_o(\mathbf{W}_o \mathbf{s}_t + \mathbf{b}_o) \end{aligned} \quad (4.26)$$

where t stands for the time-stamp and \mathbf{W}'_s is a matrix describing the weights of the feedback loop.

To train an RNN for the approximation of Bayesian experimental design, a series of experiments are simulated based on the measurement model and full Bayesian experimental design. Each simulated experiment consists of ten measurement steps of one chromatic confocal peak. The measurements and the corresponding utility functions are stored as training data for the RNN, which can be expressed in the following form:

$$\begin{aligned}
\mathbf{l}_t &= \mathbf{W}_l \lambda_t + \mathbf{b}_l \\
\mathbf{m}_t &= \mathbf{W}_m g_t + \mathbf{b}_m \\
\mathbf{s}_t &= \mathbf{l}_t \circ \mathbf{m}_t \\
\mathbf{k}_t &= \text{LSTM}(\mathbf{W}_k, \mathbf{s}_t, \mathbf{s}_{t-1}, \mathbf{b}_k) \\
\mathbf{o}_t &= \text{ReLU}(\mathbf{W}_o \mathbf{k}_t + \mathbf{b}_o)
\end{aligned} \tag{4.27}$$

where \mathbf{l}_t is a hidden layer with 200 neurons to encode the measurement location, \mathbf{m}_t is a hidden layer also with 200 neurons to encode the measured intensity. The measured wavelength at this time step is denoted λ_t and the measured signal at this time step is represented by g_t . The layer \mathbf{s}_t merges \mathbf{l}_t and \mathbf{m}_t by taking element-wise multiplication with the Hadamard operator. The layer \mathbf{k}_t is a sophisticated recurrent layer, namely the Long Short-Term Memory (LSTM) [Hoc97], which memorizes information from previous measurement steps of an experiment. The output layer is denoted by \mathbf{o}_t with the rectified linear unit (ReLU) as the activation function. The weights and biases of each layer are represented by $\mathbf{W}_{(\cdot)}$ and $\mathbf{b}_{(\cdot)}$ respectively. The collection of weight matrices for the LSTM layer is denoted by \mathbf{W}_k . The target of training is to find the weights and biases which best fits the simulated experiments and is conducted through an RMSProp optimizer with the objective of minimizing the mean squared logarithmic error. The whole process is implemented in Python based on Tensorflow [Mar15] and Keras [Cho15], and is computed using GTX 1050 graphics card by Nvidia. The training takes a couple of hours, but during measurement, the feed-forward calculation of an RNN is much faster than full Bayesian experimental design which requires multiple nested MC sampling.

As a comparison, 300 experiments of chromatic confocal measurements are simulated using three approaches: full Bayesian experimental design, approximation using RNN, and equidistant measurement. Parameters of the signal are drawn randomly. As can be seen from Figure 4.28, measurement with Bayesian experimental design has a lower mean absolute error compared with an equidistant measurement method when the number of measurements are equal. The approximation by the recurrent neural network does not perform

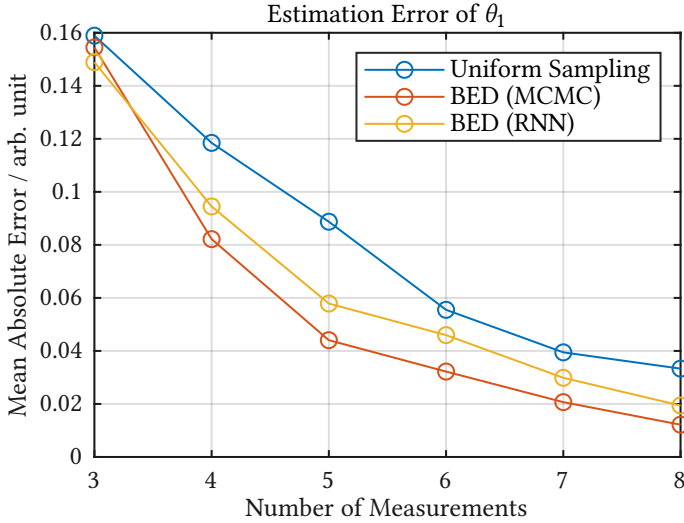


Figure 4.28: Comparison of uniform sampling, Bayesian experimental design, and Bayesian experimental design simulated through RNN.

as well as the Bayesian experimental design, due to the errors generated in the utility functions. However, it still yields a lower mean absolute error for parameter estimation compared with the equidistant measurement scheme.

Conventional feed-forward neural network with even just a single hidden layer is proven to be a universal approximator [Cyb89], which indicates that any function can be approximated by a neural network with a single hidden layer as long as the layer is large enough. The RNN is even more powerful and has been proven to be Turing-complete [Sie95]. While the training of the feed-forward neural network can be seen as optimization over functions, the training of the recurrent neural network can be seen as optimization over programs. There theoretically exists one RNN which perfectly approximates Bayesian experimental design of a specific model.

For MCMC-based BED, the sampling of the parameters with $N = 640$ and $M = 999$, which equals 640000 samples, takes roughly 18 s. With 101 number of designs, one sample of measurement result is drawn for each parameter pair

θ , which takes also roughly 18 s. The total number of measurement samples is therefore $640000 \times 101 = 64640000$. The sampling process takes a significant amount of time as the process is based on CPU (Intel i5-4210M). On the contrary, the RNN-based BED is implemented on GPU (Nvidia GTX 1050) and the complete network consists of 341901 trainable weights and biases. A single inference of all utility values for one measurement is merely 28 ms, which is more than 600 times faster than the MCMC-based BED.

4.4.4 Summary

For post-measurement refinement, a localized scan can be performed in the vicinity of the peak position based on the measurement result from the previous measurement stage. Although the scan can be implemented through uniform sampling, a more efficient dynamic sampling approach is proposed based on Bayesian experimental design. Based on the intensity measurement of several wavelength positions, the new wavelength position to be measured is determined through the computation of the utility function over possible wavelength positions. The wavelength with the largest utility value indicates the highest expected information gain if measured next. Although the computation can be implemented based on MCMC sampling, the speed of the computation is limited due to the large number of samples required. To accelerate the process, an RNN is developed and trained based on simulated experiments to approximate the computation of the utility function. The performance of the RNN-approximation is lower than full BED but higher than the uniform sampling, while the speed of the computation is 600 times faster than the MCMC-based BED.

5 Evaluation and Results

This chapter presents an evaluation of the measurement methods introduced previously within the context of the cascade measurement strategy. In Section 5.1, the configuration and characteristics of the experimental setup are presented. To evaluate the proposed measurement techniques, several benchmark measurements are performed based on conventional array confocal scanning and the results are presented in Section 5.2. Section 5.3 discusses results from the compressive shape from focus method, which is capable of locating the rough position of the object position in a wide axial range using a minimum number of frames. In Section 5.4, measurement results of the iterative array adaptation method are analyzed, where two iterations of measurements are performed. The results from the direct area scanning method are demonstrated in Section 5.5. And in Section 5.6, results from various methods are summarized and compared.

5.1 Experimental Setup

The spatial DMD, the camera sensor and the microscope arms are carefully aligned so that the camera image covers the complete effective area of the DMD. Relative planar rotation between the two coordinate systems are also minimized (ignoring the intrinsic 180° rotation) through alignment. Figure 5.1 shows an image of the 1951 USAF Resolution Test Target from Thorlabs (R1DS1P). Full-field illumination is applied by switching on all pixels on the spatial DMD. The illumination spectrum is a Gaussian function centered around 555 nm with a FWHM of 1 nm. The rectangular area which is slightly bright indicates the illumination area on the test target. According to result

of the OpticStudio simulation in Section 3.2, with diffraction limited optics, the spot size of the optical system is smaller than the pixel size of the camera even at the maximum wavelength of 680 nm. Consequently, resolution of the imaging system can be considered to be instrument limited. At 555 nm, the paraxial magnification equals 0.373, which leads to an object-side pixel size of $2.4 \mu\text{m}$. The theoretical resolution limit of the system can be calculated to be 208.3 lp mm^{-1} . As shown by the zoomed patch on the right of Figure 5.1, the system is capable of resolving test patterns until element #4 of group #7. This element indicates a resolution of 181 lp mm^{-1} , which closely matches the simulation results. Therefore, optical alignment is considered to be close to ideal.

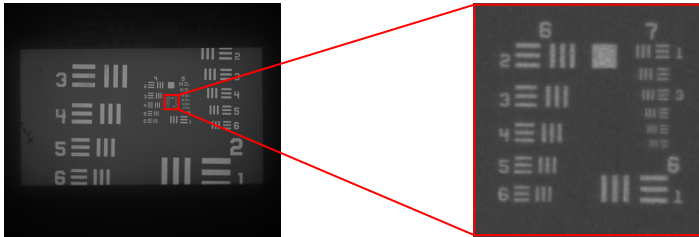


Figure 5.1: Image of 1951 USAF Resolution Test Target from Thorlabs (R1DS1P). Illumination spectrum is a Gaussian function centered around 555 nm with a FWHM of 1 nm.

A camera calibration procedure is first implemented (Section 3.2.3), after which several experiments are made based on the conventional array scanning method as a benchmark and the cascade measurement strategy. Two test targets are applied for the measurement experiments. An optical mirror is used to calibrate the image field of the optical system. Additionally, the sensitivities of various methods are characterized through a series of measurements. Secondly, a two-Euro coin is selected as a test target to demonstrate the capability of AdaScope in a practical scenario.

The system is focused on a small area on the two-Euro coin with the letters E and U. This area is chosen due to its complex structure. There are three major levels of height in this area. The bottom surface and the top surface of the letter face are both designed to be flat, whereas the middle surface of the

European map has a wavy profile packed with small indentations. All results are laterally presented in the coordinate system of the spatial DMD.



Figure 5.2: Test target for experimental investigation. (Source: from Pixabay under CC0 Creative Commons)

5.2 Benchmark: Confocal Array Scanning

To provide a benchmark for the proposed methods, conventional array scanning is implemented with AdaScope to provide an accurate but slow measurement of the target. For each pinhole array generated by the spatial DMD, a series of images are captured while the illumination wavelength is scanned.

As shown in Figure 5.3, with an axial shift of $95.25\ \mu\text{m}$ from the focal position, the blurred focal spot of a single DMD pixel easily reach a distance of 15 pixels while still maintaining sufficient energy for crosstalk. As a well optimized imaging system, the spherical aberration of the chromatic objective is aggressively corrected in order to reduce the spot size. Nevertheless, such aggressive correction often leads to a blurred spot distribution where more energy are concentrated in the outer area, making the system more susceptible to crosstalk. With the current setup, a pitch distance of 10 pixels is considered the minimum for acceptable measurement, while a pitch distance of at least 20 pixels is needed to suppress crosstalk to a minimum level. For the benchmark measurement, a pitch distance of 20 pixels has been selected, which corresponds to 400 lateral scans per axial position.

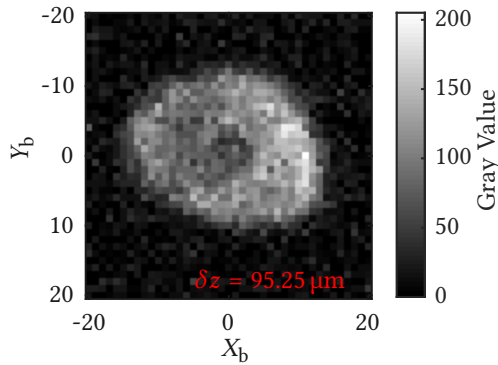


Figure 5.3: Exemplary spatial distribution of the blurred light at a distance of $95.25 \mu\text{m}$ from the focal plane, plotted in the pixel coordinate system of the spatial DMD.

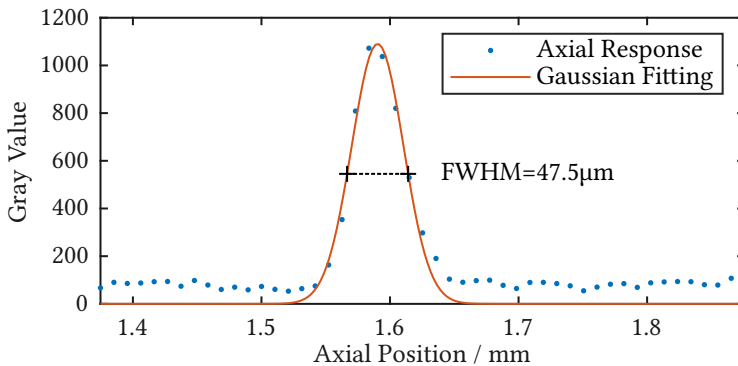


Figure 5.4: With an illumination bandwidth of 1 nm , the axial response has a FWHM of $47.5 \mu\text{m}$.

For the axial direction, the response of the system can be recorded by scanning through a series of illumination wavelengths. As illustrated in Figure 5.4, with an illumination bandwidth of 1 nm , the axial response has a FWHM of $47.5 \mu\text{m}$. To accurately locate the position of the response peak, the axial step size has to be at most half of the FWHM of the axial response, which corresponds to a number of 196 axial steps for the complete measurement range. To provide a more accurate benchmark measurement, an axial step of $10 \mu\text{m}$ has been

implemented for a smaller axial range of 0.5 mm, which is more than sufficient to cover the profile variation of the euro coin.

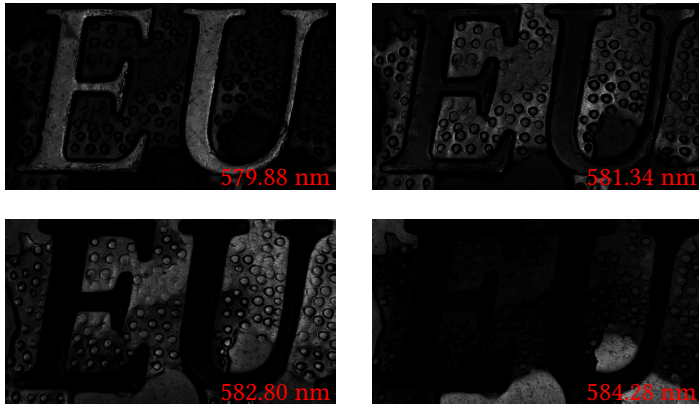


Figure 5.5: Confocal signal for the axial position #20, #23, #26 and #29. Corresponding illumination wavelengths are labeled in red.

Figure 5.5 demonstrates the measured signal at several axial positions for example. As can be seen, due to the confocal filtering, only areas which are in focus return a high intensity in the captured image, while the out of focus areas are close to black.

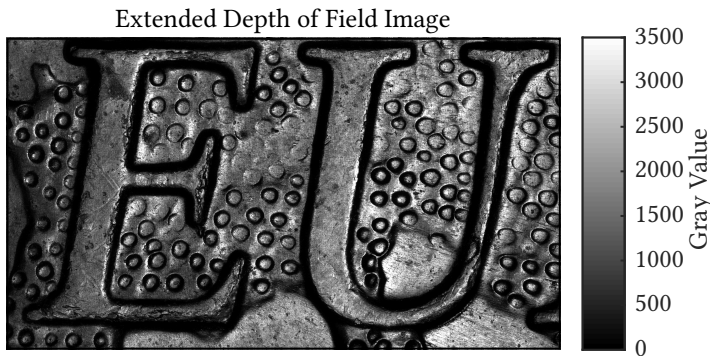


Figure 5.6: Extended depth of field image of the test target.

In post processing, by taking the maximum axial intensity for each lateral location, an extended depth of field image can be reconstructed, demonstrating the surface texture of the test target (Figure 5.6). For areas where the height changes drastically, such as the edge of the two letters, the intensity drops heavily due to self-occlusion. Confocal measurements for such areas are considered not reliable and are excluded based on a threshold set for the maximum intensity. In the following figures, such areas are illustrated with a white color.

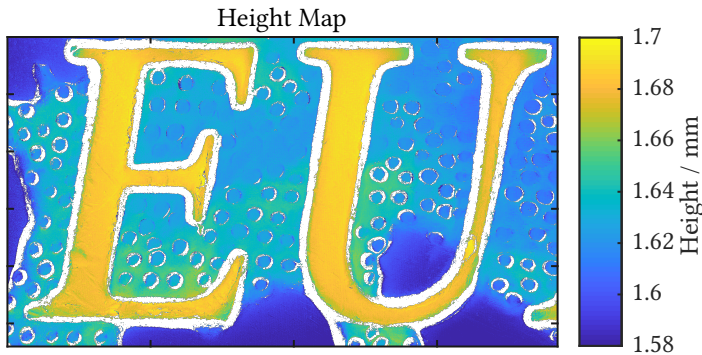


Figure 5.7: Height map reconstructed from the confocal signal of array scanning through Gaussian fitting.

To retrieve the height information of the test target, Gaussian fitting is implemented for the confocal signal, in the form of $\hat{g} = A \exp(-(z - \mu)^2 / (2\sigma^2))$ where \hat{g} represents the expectation of the gray value in the captured image and z represents the axial focus position of the corresponding wavelength. The center position of the Gaussian peak μ is considered as the height of the object, while σ is directly related with the FWHM of the signal peak. The fitting process is performed only for five data points around the maximum axial intensity where the signal to noise ratio is highest. Figure 5.7 shows the height map of the target sample reconstructed from the confocal signal with the colorbar scaled to show a height range of $120 \mu\text{m}$. As a benchmark measurement, the result of the conventional array scanning is very accurate both in the lateral directions as well as in the axial direction. The three layers of the

coin surface are faithfully reconstructed. Structural defects such as scratches can be easily located from such a height map. As a well used coin, the height difference between the top surface of the “EU” letters and the base surface of the coin is approximately 110 μm .

Despite the good accuracy and resolution, the measurement speed of the conventional array scan is relatively slow. Even for the limited axial measurement range implemented in this experiment, a total of $49 \times 20 \times 20$ scans have to be conducted, which would take more than 10 minutes to complete for a camera capable of 30 fps. For the complete measurement range with more coarse axial step, the measurement time would be at least four times longer, limiting the application of such method to only situations which are not time-critical.

5.3 Pre-measurement: Compressive Shape from Focus

In this section, measurement results of both the conventional SFF and the compressive SFF are analyzed and compared. The compressive SFF method, with its much faster measurement speed, serves as the pre-measurement in the AdaScope system, based on which the main measurement can be initialized.

As discussed in Section 4.1, conventional shape from focus method has a limited measurement speed especially in a high NA system, due to its requirement of axial focal plane sampling as densely as possible. With an NA of 0.33, the depth of field of the AdaScope is very small compared to the axial range of measurement. To make sure all relevant axial positions are covered, a series of 196 spectra are generated for the AdaScope system, each with a FWHM of 1 nm. The step size of the wavelength are chosen nonlinearly to counter the nonlinear chromatic aberration, which leads to an axial step size of 23 μm . Figure 5.8 illustrates several example frames from the focal stack with their corresponding sharpness measurement calculated using the modified Laplacian operator. As shown by the image, the sharpness measure reaches a peak when the underlying location is in focus. By tracking the axial location where

the sharpness measure reaches maximum, the surface profile of the target object can be reconstructed.

For the compressive SFF measurement, a series of measurement filters have to be constructed so that the focal stack can be acquired in a compressive manner. In the simulation presented in Section 4.1.4, such filters are constructed through PCA of various training samples based on different surface profiles and textures. In practice, the SNR of the training samples constructed from real data are not high enough generate reliable filters. Therefore, synthetic training samples are generated based on an axially shifted Gaussian signal. The first three channels of the generated filters are illustrated in Figure 5.9.

As the filters contain both positive and negative values, each filter are realized with two separate filters in practice. The filter weightings are achieved directly through the illumination spectra coupled with the axial chromatic aberration. The intensity of the illumination spectra and the camera exposure time are adjusted accordingly so that the dynamic range of the camera is fully utilized without saturation. For 14 compression filters, 28 images are captures and the final compressed frames are calculated through the subtraction between the positive images and the negative images (Figure 5.10).

After the compressed focal stack is captured, the focus measure of the full focal stack can be reconstructed with the algorithm presented in Figure 4.6. Similar to the conventional SFF method, the surface profile of the target object is reconstructed by tracking the axial location where the sharpness measure reaches maximum.

Figure 5.11 illustrates the reconstructed height maps from both the conventional SFF method and the compressive SFF method. Result from the conventional SFF method is based on a focal stack containing 196 images captured at different axial positions. The height map clearly demonstrates the base surface, the top letter surface and the wavy European flag surface in the middle. However, when examined closely, small defects occur throughout the measurement field in a random manner, mainly due to the fact that the SFF method is generally vulnerable to camera noise and the accuracy of the measurement depends strongly on the texture of the target surface. Additionally,

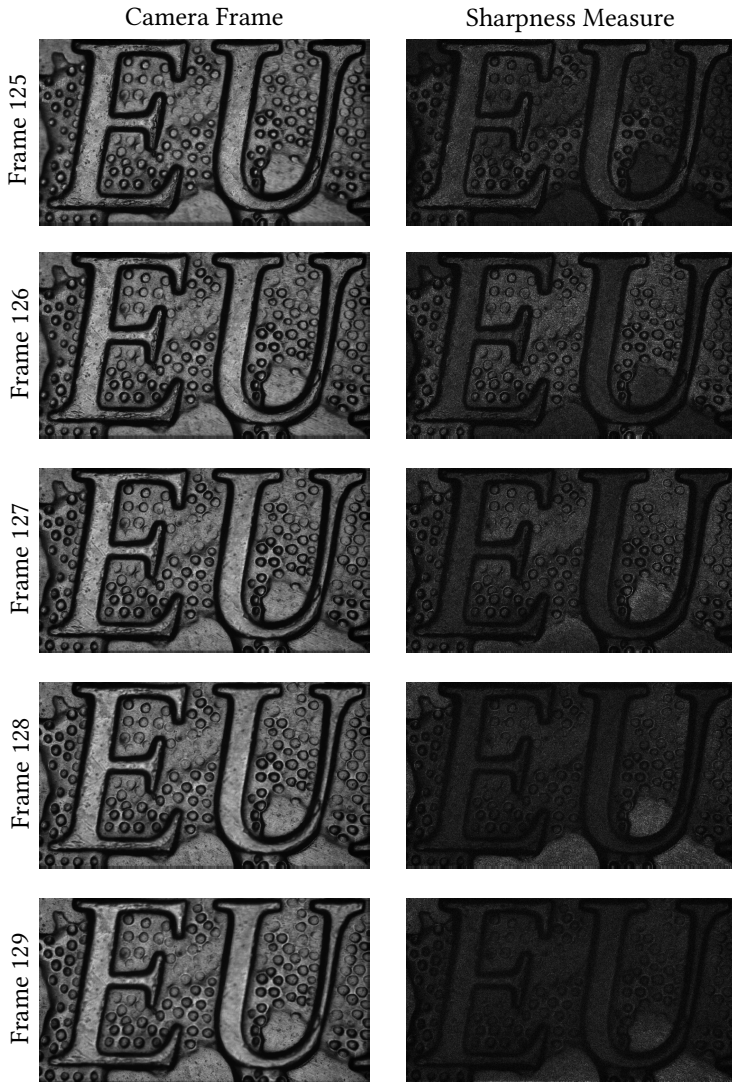


Figure 5.8: Example frames from the focal stack with their corresponding sharpness measure results. For the sharpness measure result, a brighter grey value represents a higher degree of sharpness.

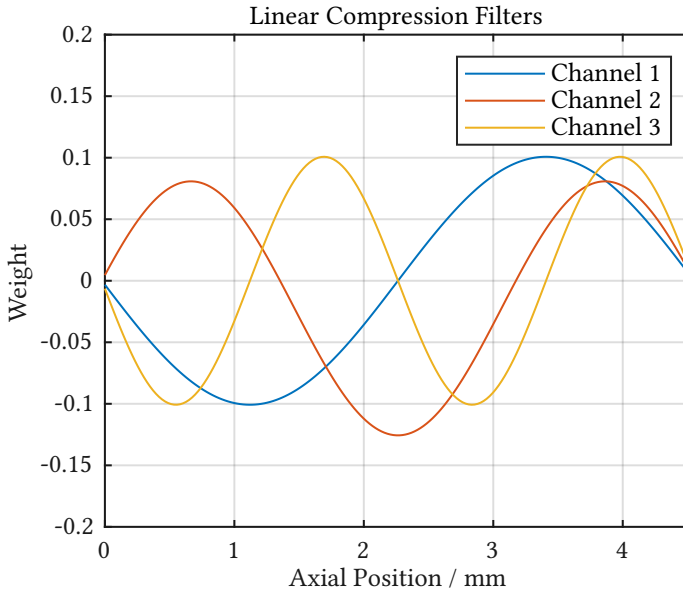


Figure 5.9: Linear compression filters of the first three channels.

the lateral smoothing procedure required when computing the focus measure also degrades the lateral resolution of the system. Overall, the conventional SFF method is considered much less reliable than the conventional confocal scan for microscopic surface profilometry, and is thus seldom applied in an industrial environment.

In the AdaScope system, instead of spending so much frame resource on conventional SFF method, which is not capable of delivering robust result, the compressive SFF method is utilized to provide a pre-measurement, based on which the main measurement stage can be initialized. With the prior knowledge that the axial region of interest is significantly smaller than the complete axial measurement range, the task of the pre-measurement range is to limit the axial measurement range in the main measurement stage so that the overall measurement efficiency can be improved. As shown in Figure 5.11, although the height map reconstructed from the compressive SFF method is

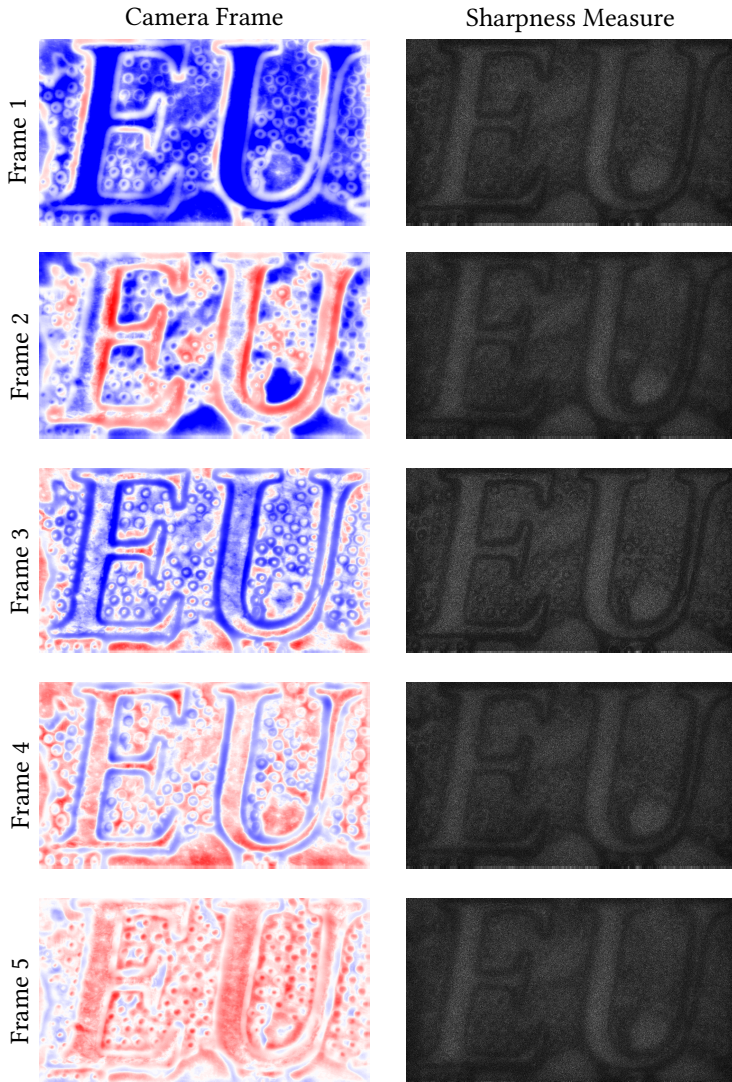


Figure 5.10: Example frames from the compressively captured focal stack with their corresponding sharpness measure results. A polar colormap is utilized as the frame contains negative values, which are indicated by the blue color. For the sharpness measure result, a brighter grey value represents a higher degree of sharpness.

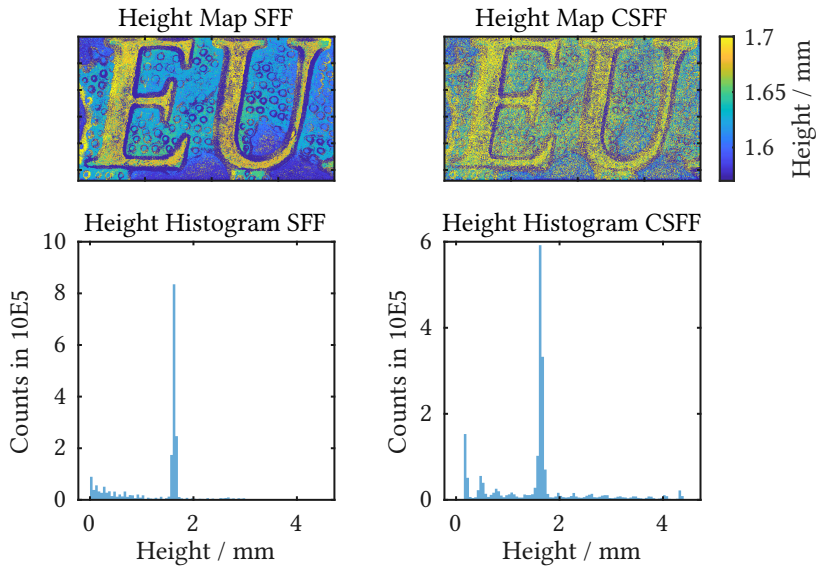


Figure 5.11: Comparison of SFF and CSFF results. Top row: height maps. Bottom row: height histogram.

more severely corrupted by noise compared to the conventional SFF result, the histogram of the height map clearly indicates that a large number of pixels are centered around 1.625 mm. By tracking the peak position of the height histogram, compressive SFF method provides a good guidance on the axial location and range for the main measurement stage with a minimum number of image captures, which is 7 times faster than the conventional SFF method.

5.4 Main measurement I: Iterative Array Adaptation

This section presents the measurement result from the iterative array adaptation method, which is considered as one candidate for the main measurement

stage. Based on the result from the compressive shape from focus measurement, the axial measurement range is limited to a length of $464\ \mu\text{m}$ centered around the axial position of $1.625\ \text{mm}$. Two iterations of the array adaptation are implemented.

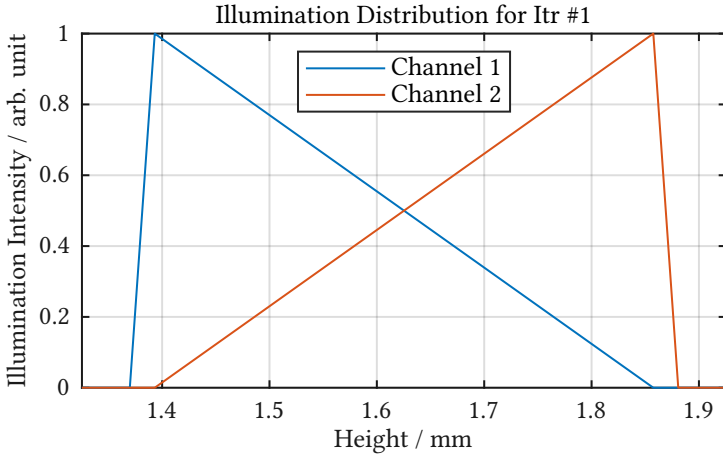


Figure 5.12: Illumination channels for the first iteration.

For the first iteration, a grid with a pitch distance of 20 pixels is scanned laterally, while the axial position is scanned through two channels of illumination spectra as shown in Figure 5.12. According to Equation 4.4, the normalized centroid position of the confocal peak, which also represents the axial position of the target, should be easily retrieved by calculating the normalized centroid position of the two channel signal. Nevertheless, due to the crosstalk between adjacent measurement locations as well as the asymmetric blurring of the out of focus light, the two centroid positions are only linearly related in a limited range.

Therefore, the relationship between of the measured signals and the axial position of the target must be calibrated (Figure 5.13). The calibration is implemented with a broadband mirror, which is mechanically scanned while the two channels of the signal are recorded. As can be seen from the calibration result, strong non-linearity starts to appear at the edges of the measurement

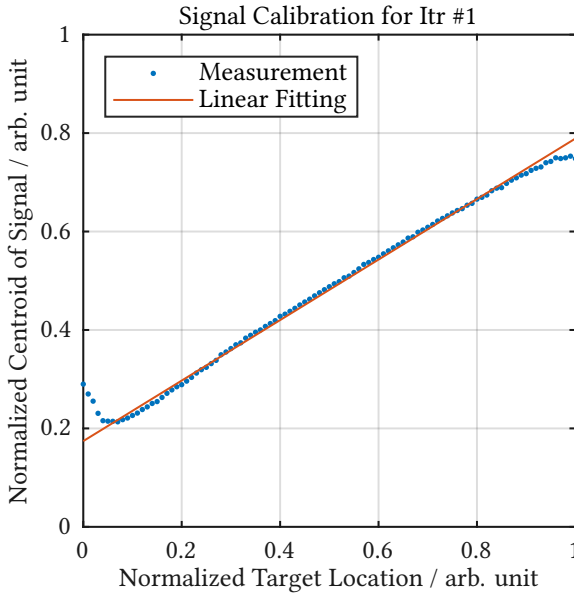


Figure 5.13: Signal calibration for iteration #1. A broadband mirror is mechanically scanned through the measurement range of iteration #1 while the signal of the two channels are recorded. The linear fitting is conducted while the 5 starting data points and 5 ending data points are excluded.

range and thus the target should be placed within the central linear region to avoid measurement defects.

Figure 5.14 shows the raw signals from iteration #1, which are reassembled from the array scanning result. Since the measurement is initialized based on the result from the compressive SFF method, the axial measurement range is chosen in a way that the target lies as close to the middle of the measurement range as possible. This results in relatively similar intensity levels in both channels. Nevertheless, the signal is sensitive enough to retrieve three dimensional information from even just two channels. For the higher regions of the “EU” letter face, the signal from channel #2 is clearly higher, while for the lower regions of the coin base surface, the signal from channel #1 is higher. By calculating the normalized centroid position of the two-channel

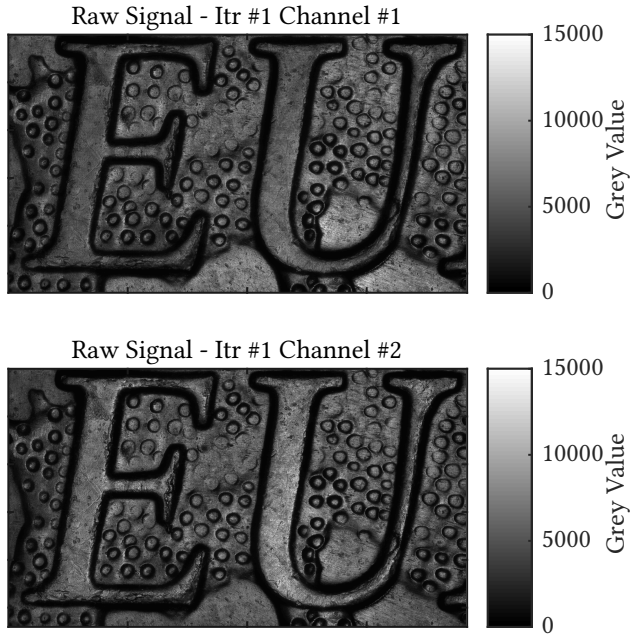


Figure 5.14: Raw signal reassembled from array scanning in iteration #1.

signal and mapping through the calibration curve, the surface profile can be reconstructed (Figure 5.15).

Even with just two channels of measurement, the reconstructed height map of the target clearly indicates three layers of structures. Nevertheless, measurement defects can also be seen across the measurement field and the mean absolute error compared to the result of array scanning from Section 5.2 is $15.3\ \mu\text{m}$.

The original array adaptation method proposed in Section 4.2 applies a binary search approach in the axial direction from iteration to iteration. Nevertheless, as shown previously, the edge of the axial measurement range often demonstrates strong non-linearity in practice, which severely affects the result of the binary search process. To counter this problem, the linear measurement filters are chosen to cover an axial range larger than the target height

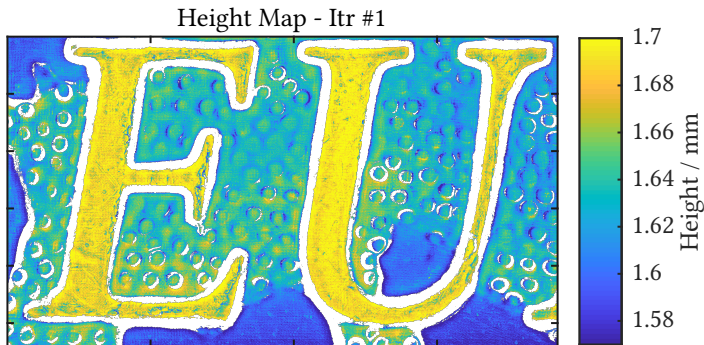


Figure 5.15: Reconstructed height map from iteration #1.

variation so that the target does not enter the edge of the measurement range. Meanwhile, in the second iteration, instead of dividing the axial measurement range from the middle, the dividing point is determined by the median value of the reconstructed height map from iteration #1. As shown in Figure 5.16, after the division, the mean value is calculated for each half based on the reconstruction result from iteration #1. The new measurement filters are centered around the two mean values while covering half of the original axial range. In this way, it is guaranteed that the target always lies as close to the center of the axial measurement range as possible. And such procedure can be implemented for further iterations as well.

In the second iteration, the lateral pitch is reduced to 10 pixels. And similar to iteration #1, signal calibration has to be implemented for the reduced pitch distance and axial measurement range. Figure 5.17 illustrates the reconstruction result from iteration #2. Although measurement artifacts still exist, the mean absolute error with respect to the confocal array scanning has dropped to $14.7 \mu\text{m}$.

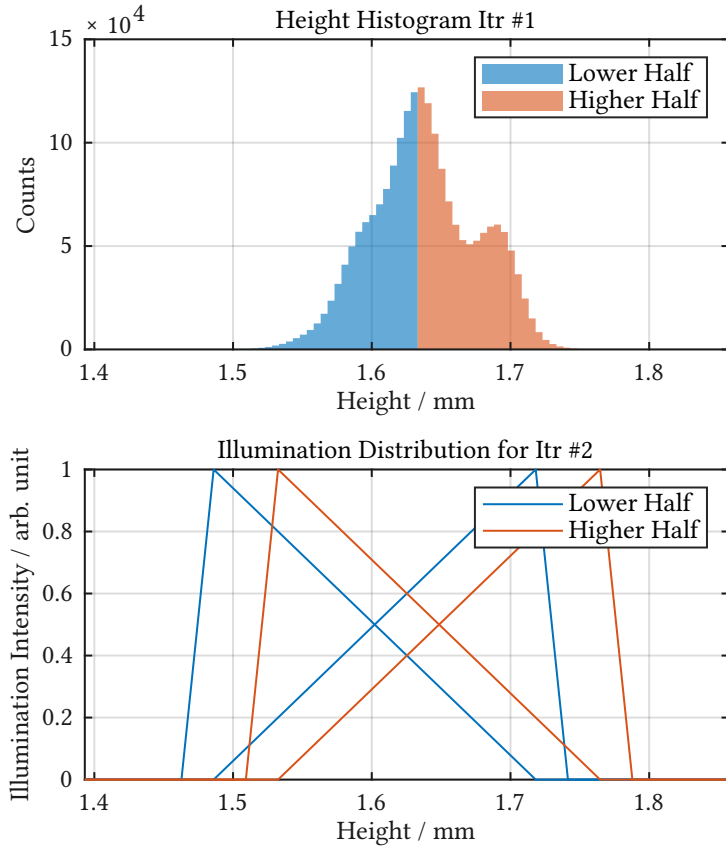


Figure 5.16: The histogram of the height reconstructed from iteration #1 can be divided into two halves which are measured by different linear filters in iteration #2. The span of the x -axes in both figures represents the axial range of iteration #1.

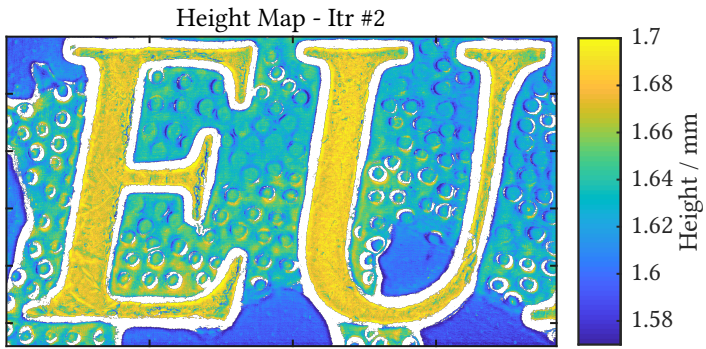


Figure 5.17: Reconstructed height map from iteration #2.

5.5 Main measurement II: Direct Area Scanning

This section presents the measurement result from the direct area confocal scanning method, which is considered as another candidate for the main measurement stage. Based on the result from the compressive shape from focus measurement, the axial measurement range is limited to a length of $501\ \mu\text{m}$, which is centered around $1.625\ \text{mm}$. The range is slightly different from the axial range used in the array adaptation method to guarantee the optimum tilting angle of the illumination field according to the lateral pitch of the spatial DMD and the magnification of the imaging system.



Figure 5.18: Left column: raw camera frames from direct area scanning. Right column: confocal signal through reordering of the raw signal.

Figure 5.18 illustrates the captured signal while the periodic tilted illumination field is scanned laterally. The left column shows exemplary raw frames from the camera, while the right column shows images in which the signal are reordered so that each frame represents the signal of a certain wavelength. Compared to the confocal signal from the conventional array scanning method shown in Figure 5.5, the background level is clearly higher, due to crosstalk from the adjacent locations in such a wide-field setup. Nevertheless, with the help of the tilting angle, a confocal peak is still visible as different areas get brighter when imaged in focus.

Similar to the post processing for conventional array scanning, Gaussian fitting is also implemented for the area scanning signal in a window centered around the maximum intensity. Figure 5.19 illustrates the measurement result using the proposed area scanning method, where several differences can be demonstrated.

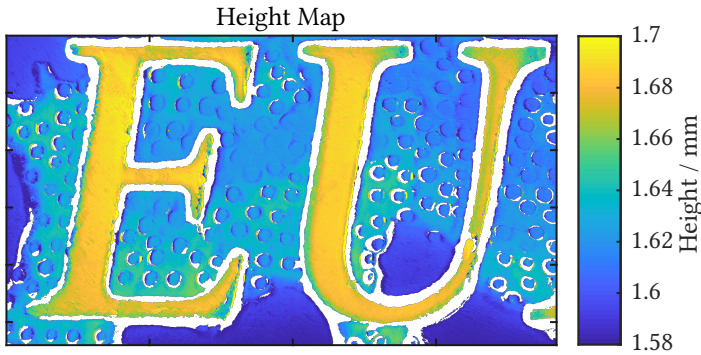


Figure 5.19: Height map reconstructed from the confocal signal of area scanning through Gaussian fitting.

Compared to the measurement result of iterative array adaptation (Figure 5.15), the measurement result is clearly much smoother and very close to the benchmark measurement by the array scanning method (Figure 5.7). As a quantitative comparison, the mean absolute error with respect to the benchmark measurement has further dropped to $8.9\mu\text{m}$. Despite these differences, the measurement result using the proposed method is completely

usable, revealing the three structure layers of the coin accurately. Center areas of the indentation holes can be measured without being much affected by the shadows around. The wavy profile of the European map is also truthfully recovered.

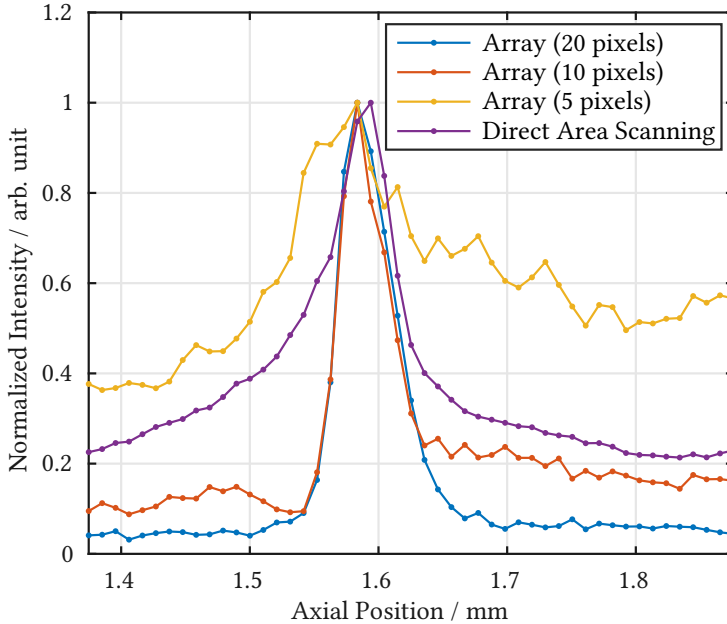


Figure 5.20: Comparison of confocal signals from array scanning and direct area scanning.

A comparison of the confocal signal from the array scanning method and the confocal signal from the direct area scanning method is presented in Figure 5.20. All signals are extracted from the same location on the coin in a relatively flat area which is located at the top left corner of the FoV. For the array scanning method, as the pitch increases, the background noise arises, leading to a wider confocal peak. The signal shape of the direct area scan method lies between the array scan methods with 10 pixels and 5 pixels.

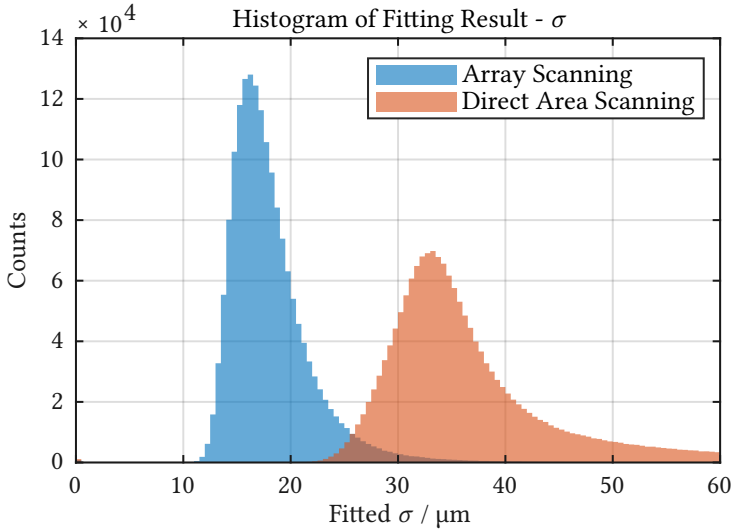


Figure 5.21: Histogram comparison of the fitting result for σ . For the array scanning method, the pitch distance is 20 pixels.

Figure 5.21 illustrates the histograms of the fitting result of σ for all measurement positions on the coin. The FWHM of the fitted Gaussian peak can be calculated with $\text{FWHM} = 2\sqrt{2\ln 2}\sigma \approx 2.35\sigma$. For the conventional array scanning method with a pitch distance of 20 pixels, the fitted σ values have an average of $18.2\ \mu\text{m}$, i.e. $42.7\ \mu\text{m}$ in terms of FWHM. For direct area scanning, the mean fitted σ value equals to $39.1\ \mu\text{m}$, which corresponds to a FWHM of $91.9\ \mu\text{m}$. In both cases, the measured axial responses are wider than the theoretical analysis. Multiple factors are responsible for this effect. First and foremost, the illumination Gaussian spectra have a certain bandwidth instead of being a Dirac pulse. The axial spanning due to the chromatic aberration for the selected spectral bandwidth (1 nm) in the current measurement range alone amounts to approximately $20\ \mu\text{m}$. Secondly, various optical aberrations in the practical system, the calibration error of the camera as well as camera noises all contribute to the broadening of the intensity peak. Thirdly, the lateral extension of the spatial DMD pixel also increases the illumination spot

size which leads to a wider axial FWHM. Additionally, the periodic illumination pattern used in the direct area scanning method also contributes to the cross-talk, which raises the background level of the signal, leading to a wider peak.

5.6 Analysis and Comparison

In this section, the accuracy of the two main measurement methods are compared against the array scanning method with fixed pitch. As trivial as it may seem, the characterization of the accuracy of a measurement system is a very complex task. The errors of the measurement can be categorized into two classes: random error and systematic error. Although the random error can often be characterized in a more straightforward manner through statistical analysis, the systematic error is much more challenging to quantify and could vary from object to object. In practice, the combination of both contributes to the final measurement error, making it difficult to separate one factor from the other. While the camera noise and the instability of the illumination source contribute most to the random error, the systematic error mostly originates from the crosstalk between the adjacent measurement locations.

To test their robustness against the random error, a series of experiments are conducted for each method with a broadband mirror as the test target. The mirror is chosen as the target due to its simple structure, which helps to suppress and simplify the effect of crosstalk. A total of 5184 lateral locations uniformly spread across the measurement field are observed through 25 repeated measurements. The standard deviations of these measurements are used as an indicator of the random part of the measurement uncertainty. The minimum, average and maximum standard deviation values of all lateral locations are listed in Table 5.1.

Several insights can be gained through this table. First of all, for the array scan method with fixed pitch, the measurement uncertainty generally increases as the pitch distance decreases. The influence of the crosstalk can be clearly demonstrated from these three different configurations. The level of crosstalk

Table 5.1: Uncertainty and speed comparison of various methods. The results are for a broadband mirror as the test target.

Method	Uncertainty			Frames per Measurement
	min	avg	max	
Array scan with fixed pitch				
Pitch: 20 pixels	0.16 μm	0.49 μm	1.37 μm	19600
Pitch: 10 pixels	0.15 μm	0.53 μm	2.12 μm	4900
Pitch: 5 pixels	0.15 μm	0.57 μm	5.69 μm	1225
Array scan with iterative array adaptation				
Iteration 1	2.21 μm	4.65 μm	8.89 μm	800
Iteration 1+2	1.31 μm	3.12 μm	6.00 μm	1000
Direct area scan				
Period: 49 pixels	0.19 μm	0.68 μm	2.39 μm	49

is not uniform across the measurement field even for a mirror because the PSF varies laterally due to optical aberrations in the system. The minimum uncertainty values remain almost unchanged, since it represents the best case in all three configurations where the crosstalk level is very small. In such situations, the uncertainty values come from the measurement of a pure confocal peak signal with minimum crosstalk from its neighbors. On the contrary, the maximum uncertainty values vary dramatically, as they represent the worse case scenario for the three configurations, in which the crosstalk level is directly linked with the pitch distance. Secondly, the direct area scan method is clearly better than the iterative array adaptation method in terms of uncertainty and has a performance close to the area scan method with a pitch distance of 10 pixels.

The data listed in Table 5.1 is not sufficient to generate a conclusion regarding the accuracy of the various methods since it does not take all systematic errors into consideration. For example, the uncertainty values in the table could lead to a false impression that the iterative array adaptation method has the worst performance of all listed methods, which is not true. Figure 5.22 illustrates the measurement result of the coin using array scan method with a pitch distance

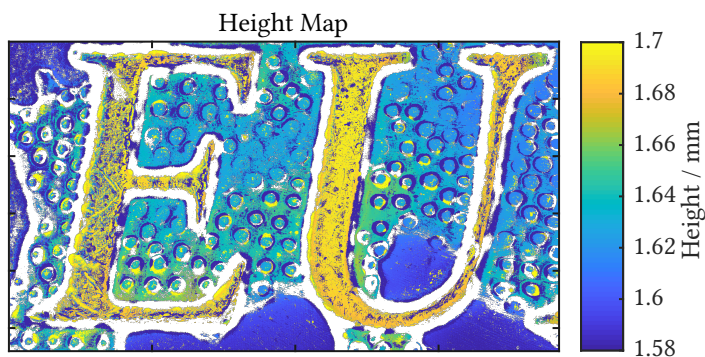


Figure 5.22: Height map reconstructed from the confocal signal of array scanning with a pitch distance of 5 pixels through Gaussian fitting. Errors in measurement due to signal crosstalk are easily visible.

of 5 pixels. The reconstructed height map contains a large amount of systematic error due to crosstalk, which makes the measurement quality much inferior to result of the iterative array adaptation method (Figure 5.17).

Table 5.2: Mean absolute bias of various methods with respect to confocal array scanning with a pitch distance of 20 pixels. The test target is the coin.

Methods	Mean Absolute Bias	Frames per Measurement
Array scan with fixed pitch		
Pitch: 20 pixels	N/A	19600
Pitch: 10 pixels	20.9 μm	4900
Pitch: 5 pixels	49.7 μm	1225
Array scan with iterative array adaptation		
Iteration 1	15.3 μm	800
Iteration 1+2	14.7 μm	1000
Direct Area Scan		
Period: 49 pixels	8.9 μm	49

The measurement of array scan with a pitch distance of 20 pixels generates the best result due to its low uncertainty and crosstalk, which has been used as a benchmark in previous sections. With this measurement result as the

reference, the mean absolute biases of other methods and configurations can be computed, which are listed in Table 5.2. Both main measurement method candidates generate a lower bias compared to array scan with a pitch distance larger than 10 pixels, indicating a smaller amount of crosstalk. In particular, the direct area scan method yields a result closest to the array scanning method with a pitch of 20 pixels.

Combining both the uncertainty and the bias, it is clear that the direct area scan method is the better one from the two candidate methods for the main measurement. With the lowest bias and a very low level of uncertainty, the performance of the direct area scan method is at least as good as the array scan method with a pitch of 10 pixels. With just 49 frames per measurement for the limited measurement range, it achieves a speed improvement of at least 100 times while maintaining a comparable measurement accuracy.

6 Concluding Remarks

The work presented in this dissertation aims to achieve high-speed surface profilometry based on an adaptive microscope with axial chromatic encoding. In this closing chapter, the results of the work are summarized and an outlook is presented for future research topics.

6.1 Conclusion

Conventional optical surface profilometry technology is dominated by the dilemma between measurement speed and accuracy. Shape from focus methods are very efficient in terms of information acquisition but is restricted by a low accuracy in both the lateral and the axial directions. Despite the high resolution and accuracy of the confocal systems, the measurement suffers from a slow speed due to its dependence on scanning. Even with state-of-the-art array scanning methods, a minimum pitch between adjacent measurement locations must be maintained to avoid crosstalk. To tackle this problem, a holistic approach has been taken to design and develop an adaptive microscope, i.e., the AdaScope, together with a cascade measurement strategy.

The AdaScope is composed of two subsystems. To begin with, the programmable light source is developed based on a supercontinuum laser. A cross disperser pair of a dispersion prism and an echelle grating has been constructed to generate the 2D dispersion pattern of the laser spectrum, i.e., its echellogram, which is projected onto a DMD. By switching the micro mirrors on the DMD individually, arbitrary spectra can be generated and collected by the liquid light guide for output to the next stage. Through careful alignment and calibration, the programmable light source is able

to generate Gaussian peak spectra with a FWHM smaller than 1 nm. And when used as a scanning source, a step size as accurate as 0.01 nm can be achieved. The second subsystem is a programmable array microscope based on a second DMD. Light from the programmable light source is homogenized and projected onto the second DMD for spatial filtering. Each pixel on the second DMD serves as a secondary light source which can be individually addressed. A chromatic objective is applied in the microscope to achieve axial chromatic encoding. Through combination of the illumination spectra and the spatial DMD pattern, any location within the 3D measurement volume can be addressed through a localized illumination field. The combination of the two subsystems forms a flexible and adaptive microscopic system, allowing different measurement principles to be implemented and analyzed.

Based on the AdaScope platform, a cascade measurement strategy has been proposed. Multiple measurement methods can be combined to perform a complete measurement task, where raw and fast measurement result in one stage is used to initialize slower but more accurate measurement in the next stage. By combining the advantages of different methods, the dilemma between scanning density and measurement accuracy in conventional optical surface profilometry can be tackled. For the pre-measurement stage, a compressive SFF method has been developed to generate raw measurement result of the surface profile using a small number of frames. Each frame is compressively captured as a linear combination of all focal planes within the measurement volume. Reconstruction of the SFF signal is directly performed in the focus measure space. Experimental results demonstrate an acquisition speed 7 times faster than conventional SFF when an estimation of the object axial position can still be correctly performed. As one candidate method for the main measurement stage, the iterative array adaptation method is based on the conventional confocal array scanning method. Multiple iterations of array scanning are performed while the array density and the axial measurement range are dynamically adjusted from iteration to iteration. The mean absolute error with respect to benchmark measurement performed by the conventional array scanning method is 14.7 μm . Another candidate method for the main measurement stage is the direct area scanning method based on a tilted illumination field. It has been demonstrated both theoretically and experimentally

that when an area illumination field is tilted to a specific angle according to the NA of the system, the confocal signal, which would disappear in a wide-field microscope, can be largely preserved, providing a cue for the 3D surface profile. Compared to the conventional array scanning method with a pitch of 20 pixels, the axial FWHM of the direct area scanning method is doubled, indicating a moderately reduced sensitivity. Nevertheless, compared to array scanning with a pitch of 10 pixels, the measurement speed is more than 100 times faster, while achieving a comparable measurement accuracy. Last but not least, for post measurement refinement, localized confocal scanning based on Bayesian Experimental Design has been discussed. Conventional Bayesian Experimental Design is computationally intensity due to the requirement of a nested MCMC sampling when calculating the utility function. Simulation results has demonstrated that this process can be much accelerated through the approximation by an RNN, which is able to achieve performance between uniform sampling and BED-enabled sampling. Compared to the full BED, the calculation of the utility function is accelerated by a factor of 600 by the RNN-based BED.

Through the combination of these different methods, the information regarding the 3D surface profile of the target object can be acquired in a much more efficient way. Due to the intrinsic adaptability of the AdaScope, the prior information regarding the test target, such as the CAD model, can be easily incorporated into the measurement process to further accelerate the measurement speed. Additionally, the measurement range of the system can be dynamically adjusted for individual applications. These characteristics grant the AdaScope the unique capability to swiftly adapt to different inspection tasks, which matches the need of a smart factory in the era of Industry 4.0.

6.2 Outlook

Although impressive results have been reported based on the AdaScope platform, its performance is still bounded by several limitations. As an outlook, potential improvements in future research are identified and discussed.

Incorporation of New Measurement Mode

First of all, thanks to the adaptability of the AdaScope system, new measurement principles can be easily implemented and incorporated. One particularly interesting area of research is chromatic confocal spectral interferometry (CCSI). The CCSI technology is first proposed by Papastathopoulos et al. [Pap06] as a novel method for topography measurements, which combines the techniques of spectral interferometry and chromatic confocal microscopy. It is the first interferometric method that utilizes a confocally filtered and chromatically dispersed focus for detection. Unlike white light interferometry, the depth range of the sensor is decoupled from the NA of the microscope objective with the chromatically dispersed focus. As an interferometric method, the measured signal provides an even higher sensitivity compared to the confocal measurement. The beam trap in the current AdaScope setup (Figure 3.21) can be replaced by a switchable reference arm with phase compensation. This would allow a final stage of interferometric measurement to be incorporated into the cascade measurement strategy.

Better Illumination

Secondly, the illumination system can be further improved. This is critical to the AdaScope system as the quality of the illumination directly affects the measurement accuracy. On one hand, the spectral accuracy of the illumination generation can be increased through a new calibration procedure. Currently, the spectral responses are captured for a series of scanning macro pixels for calibration. Due to the small intensity of the light on a single macro pixel, the recorded response suffers from a low SNR mainly due to the photon noise. Kang et al. [Kan18] proposed a novel calibration procedure for DMD-enabled programmable optical filter based on Hadamard transform. Instead of scanning individual pixels, patterns generated from Hadamard transform are projected onto the DMD while the respective spectral responses are recorded. With much higher intensity of the reflected light, the spectrum measurement benefits from a higher SNR. Compared to sequential scanning, spectra generated through calibration based on Hadamard transform are more accurate, particularly when the number of channels (pixels) are higher. Although the proposed method is only applied in a 1D DMD chip with a relatively small

number of channels, the principle applies to a 2D DMD as well. To adopt this calibration procedure in the AdaScope system, modifications to the algorithm has to be developed to account for the higher requirement of computation due to the larger number of pixels. Apart from the spectral accuracy, the spatial homogeneity of the light projected onto the spatial DMD is also of great importance, particularly for methods using complex illumination spectra, such as the CSFF method and the iterative array adaptation method. To improve the homogeneity of the illumination field, customized microlens arrays with a smaller lenslet pitch and a higher number of periods can be applied.

The superior adaptability of the AdaScope system originates from its ability to generate different 3D illumination fields. Nevertheless, time-multiplexing techniques are highly involved in the current system architecture. Although not intrinsically required by any methods proposed in this dissertation, the time-multiplexing procedure directly limits practical performance of the AdaScope. The homogenized light from the programmable light source is projected to the complete spatial DMD, where switchable pixels serve as secondary light sources. Under such circumstances, illumination spectra of different lateral positions can not be adjusted simultaneously. As discussed in Section 3.2.4, the control of the illumination is split to the control of the temporal illumination spectrum and the control of the temporal spatial DMD pattern. For the measurement methods proposed in this dissertation, one of the aforementioned components is fixed empirically while the other component is derived from the desired complete illumination field. For example, in the direct area scanning method, a series of scanning wavelength peaks is fixed before the corresponding spatial DMD patterns are derived. Although efficient enough for most methods where a regular (e.g., symmetric, periodic, global, etc.) illumination field is required, this empirical procedure becomes very inefficient when the required illumination field has a more complex spatial distribution. Thus more rigorous algorithms should be developed in the future for the decomposition of the illumination matrix/tensor, with the target of minimizing the total time of illumination. Toward a more distant future, hyperspectral displays can be built based on technologies such as quantum dot, enabling simultaneous adjustment of the illumination spectra for different lateral locations.

Implementation of Localized Scan

In Chapter 5, only the pre-measurement stage and the main measurement stage of the cascade measurement strategy are evaluated through practical experiments. For post-measurement refinement, the idea of dynamic localized scan based on Bayesian Experimental Design has been discussed in Section 4.4. Nevertheless, this method is only analyzed through simulation for two reasons. On one hand, as discussed previously, the current illumination generation scheme relies heavily on time-multiplexing, which makes the localized scan very inefficient. To have a meaningful implementation of any localized scanning method, a dramatically different hardware configuration has to be designed and developed, allowing simultaneous spectral controls of individual lateral locations. On the other hand, despite the great speed improvement of RNN-based BED compared to full BED based on MC sampling demonstrated by the simulation, the inference of the utility function over a relatively small number of designs for a single lateral location still takes 0.028 s. For the full HD resolution of the spatial DMD, the inference would take more than one hour to complete. This problem can be partly circumvented by the next generation GPU technology, such as RTX 2080 by Nvidia. More importantly, new neural network architectures remain to be investigated to improve the performance of the proposed method. For example, parallelization of multiple lateral locations has a high potential considering the increasing RAM of the graphic card. Meanwhile, information from adjacent locations can also be incorporated to increase the accuracy of the inference.

A Appendix

A.1 Camera Calibration

To target of the camera calibration process is to get the correspondence between the spatial DMD coordinate system and the camera coordinate system, so that the dynamic control of the DMD can be derived based on camera images. The process is based on a pinhole camera model as presented in the work by Zhang [Zha00] and implemented using the OpenCV library in Python.

Firstly, a single wavelength illumination is projected onto the spatial DMD and a broadband mirror perpendicular to the optical axis is positioned axially so that the image of the DMD pattern is sharply focused on the camera image. Secondly, the spatial DMD displays a point array pattern with a pitch distance of 20 pixels. Lastly, the image of the point array is captured by the camera.

The coordinates of the point array in the camera image is retrieved by calculating the intensity centroid position inside a region of interest with a width of 15 camera pixels centered around the brightest pixel for each point. With the list of DMD coordinates and the list of camera coordinates for the point array, the rotation vector, the translation vector, the camera matrix as well as distortion coefficients are calculated based on existing method from the OpenCV library.

Based on the calibration result, the DMD pixels can be projected onto the camera plane with their corresponding intensities interpolated from the camera image.

A.2 Bernstein Polynomials

For the measurement method based on linearly weighted filters, a measurement matrix constructed with the Bernstein polynomials has the property of maintaining the normalized centroid position of the underlying signal. The proof is given in this section.

The $n + 1$ Bernstein basis polynomials of degree n are defined as

$$b_{v,n}(t) = \binom{n}{v} t^v (1-t)^{n-v} \quad (\text{A.1})$$

$$v = 0, \dots, n \quad t \in [0,1].$$

The signal is represented by a vector $\mathbf{x} = (x_1, \dots, x_o)^\top$ and the measurement can be represented by $\mathbf{y} = (y_0, \dots, y_n)^\top$ with $n + 1$ channels. Suppose both the signal and the measurement are spanned on a support of $[0,1]$, the normalized centroid of the signal and the measurement can be expressed as

$$\text{COG}(\mathbf{x}) = \frac{\sum_{i=1}^o \frac{i-1}{o-1} x_i}{\sum_{i=1}^o x_i}, \quad (\text{A.2})$$

$$\text{COG}(\mathbf{y}) = \frac{\sum_{v=0}^n \frac{v}{n} y_v}{\sum_{v=0}^n y_v}.$$

If each channel of the measurement is based on the respective Bernstein polynomial, the result can be expressed as a linear combination of the signal components

$$y_v = \sum_{i=1}^o b_{v,n}\left(\frac{i-1}{o-1}\right) x_i. \quad (\text{A.3})$$

Therefore, the centroid of the measurement can be rewritten as

$$\begin{aligned}
\text{COG}(y) &= \frac{\sum_{v=0}^n \frac{v}{n} y_v}{\sum_{v=0}^n y_v} \\
&= \frac{\sum_{v=0}^n \frac{v}{n} \sum_{i=1}^o b_{v,n}\left(\frac{i-1}{o-1}\right) x_i}{\sum_{v=0}^n \sum_{i=1}^o b_{v,n}\left(\frac{i-1}{o-1}\right) x_i} \\
&= \frac{\sum_{i=1}^o x_i \sum_{v=0}^n \frac{v}{n} b_{v,n}\left(\frac{i-1}{o-1}\right)}{\sum_{i=1}^o x_i \sum_{v=0}^n b_{v,n}\left(\frac{i-1}{o-1}\right)}.
\end{aligned} \tag{A.4}$$

All Bernstein polynomials of a certain degree forms a partition of unity:

$$\sum_{v=0}^n b_{v,n}(t) = \frac{n!}{v!(n-v)!} t^v (1-t)^{n-v} = [t + (1-t)]^n = 1. \tag{A.5}$$

Meanwhile, it can be shown that

$$\begin{aligned}
\sum_{v=0}^n \frac{v}{n} b_{v,n}(t) &= \sum_{v=1}^n \frac{v}{n} b_{v,n}(t) \\
&= \sum_{v=1}^n \frac{v}{n} \frac{n!}{v!(n-v)!} t^v (1-t)^{n-v} \\
&= t \sum_{v=1}^n \frac{(n-1)!}{(v-1)! [(n-1)-(v-1)]!} t^{v-1} (1-t)^{[(n-1)-(v-1)]} \tag{A.6} \\
&= t \sum_{v=0}^{n-1} b_{v,n-1}(t) \\
&= t.
\end{aligned}$$

Based on Equations A.4, A.5 and A.6, it can be proven that the normalized centroid of the measurement equals that the of signal:

$$\begin{aligned}\text{COG}(\mathbf{y}) &= \frac{\sum_{i=1}^o x_i \sum_{v=0}^n \frac{v}{n} b_{v,n}\left(\frac{i-1}{o-1}\right)}{\sum_{i=1}^o x_i \sum_{v=0}^n b_{v,n}\left(\frac{i-1}{o-1}\right)} \\ &= \frac{\sum_{i=1}^o x_i \frac{i-1}{o-1}}{\sum_{i=1}^o x_i} \\ &= \text{COG}(\mathbf{x}).\end{aligned}\tag{A.7}$$

Bibliography

- [Arr10] ARRASMITH, Christopher L.; DICKENSHEETS, David L. and MAHADEVAN-JANSEN, Anita: “MEMS-based handheld confocal microscope for in-vivo skin imaging”. In: *Optics Express* 18.4 (Feb. 2010), p. 3805. DOI: 10.1364/OE.18.003805.
- [Asi01] ASIF, M. and CHOI, Tae-Sun: “Shape from focus using multilayer feedforward neural networks”. In: *Image Processing, IEEE Transactions on* 10.11 (Nov. 2001), pp. 1670–1675. DOI: 10.1109/83.967395.
- [Ba14] BA, Jimmy; MNIH, Volodymyr and KAVUKCUOGLU, Koray: “Multiple Object Recognition with Visual Attention”. In: *arXiv e-prints* (Dec. 2014).
- [Bar79] BARTELT, H. O.: “Transmission of two-dimensional images by wavelength multiplexing”. In: *Optics Communications* 28.1 (Jan. 1979), pp. 45–50. DOI: 10.1016/0030-4018(79)90166-4.
- [Bei04] BEI, Ling; DENNIS, Glenn I.; MILLER, Heather M.; SPAINE, Thomas W. and CARNAHAN, Jon W.: “Acousto-optic tunable filters: fundamentals and applications as applied to chemical analysis techniques”. In: *Progress in Quantum Electronics* 28.2 (2004), pp. 67–87. DOI: 10.1016/S0079-6727(03)00083-1.
- [Bou05] BOUDOUX, C.; YUN, S. H.; OH, W. Y.; WHITE, W. M.; IF-TIMIA, N. V.; SHISHKOV, M.; BOUMA, B. E. and TEARNEY, G. J.: “Rapid wavelength-swept spectrally encoded confocal microscopy”. In: *Optics Express* 13.20 (Oct. 2005), p. 8214. DOI: 10.1364/OPEX.13.008214.
- [Bra00] BRADSKI, G.: “The OpenCV Library”. In: *Dr. Dobb’s Journal of Software Tools* (2000).

- [Bue02] BUETTNER, Alexander and ZEITNER, Uwe D.: “Wave optical analysis of light-emitting diode beam shaping using microlens arrays”. In: *Optical Engineering* 41 (2002), p. 9. DOI: 10.1117/1.1499975.
- [Can06] CANDÈS, Emmanuel J.; ROMBERG, Justin K. and TAO, Terence: “Stable signal recovery from incomplete and inaccurate measurements”. In: *Communications on Pure and Applied Mathematics* 59.8 (2006), pp. 1207–1223. DOI: 10.1002/cpa.20124.
- [Cha00] CHA, Sungdo; LIN, Paul C.; ZHU, Lijun; SUN, Pang-Chen and FAINMAN, Yeshaiahu: “Nontranslational three-dimensional profilometry by chromatic confocal microscopy with dynamically configurable micromirror scanning”. In: *Appl. Opt.* 39.16 (June 2000), pp. 2605–2613. DOI: 10.1364/AO.39.002605.
- [Cha15] CHAKROVA, Nadya; RIEGER, Bernd and STALLINGA, Sjoerd: “Development of a DMD-based fluorescence microscope”. In: *Proc. SPIE* 9330 (2015). DOI: 10.1117/12.2077677.
- [Cha17] CHANG, Ting-Jui and SU, Guo-Dung J.: “A confocal microscope with programmable aperture arrays by polymer–dispersed liquid crystal”. In: *Proc. SPIE* 10376 (2017). DOI: 10.1117/12.2273887.
- [Cha95] CHALONER, Kathryn and VERDINELLI, Isabella: “Bayesian Experimental Design: A Review”. In: *Statist. Sci.* 10.3 (Aug. 1995), pp. 273–304. DOI: 10.1214/ss/1177009939.
- [Che11] CHEN, Liang-Chia; LI, Hau-Wei and CHANG, Yi-Wei: “Full-field chromatic confocal surface profilometry employing DMD correspondence for minimizing lateral cross talks”. In: *Proc. SPIE* 8321 (2011), p. 8. DOI: 10.1117/12.904881.
- [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: *Optics Express* 21.20 (Oct. 2013), pp. 23611–23618. DOI: 10.1364/OE.21.023611.
- [Cho15] CHOLLET, François et al.: Keras. <https://github.com/fchollet/keras>. 2015.

- [Chu06] CHUANG, C. H. and LO, Y. L.: “Digital programmable light spectrum synthesis system using a digital micromirror device”. In: *Appl. Opt.* 45.32 (Nov. 2006), pp. 8308–8314. DOI: 10.1364/AO.45.008308.
- [Cyb89] CYBENKO, G.: “Approximation by superpositions of a sigmoidal function”. In: *Mathematics of Control, Signals and Systems* 2.4 (1989), pp. 303–314. DOI: 10.1007/BF02551274.
- [Din11] DING, Yunhong; PU, Minhao; LIU, Liu; XU, Jing; PEUCHERET, Christophe; ZHANG, Xinliang; HUANG, Dexiu and OU, Haiyan: “Bandwidth and wavelength-tunable optical bandpass filter based on silicon microring-MZI structure”. In: *Opt. Express* 19.7 (Mar. 2011), pp. 6462–6470. DOI: 10.1364/OE.19.006462.
- [Dob97] DOBSON, Sarah L.; SUN, Pang-chen and FAINMAN, Yeshayahu: “Diffractive lenses for chromatic confocal imaging”. In: *Appl. Opt.* 36.20 (July 1997), pp. 4744–4748. DOI: 10.1364/AO.36.004744.
- [Duo14] DUCASTELLA, Martí; VICIDOMINI, Giuseppe and DIASPRO, Alberto: “Simultaneous multiplane confocal microscopy using acoustic tunable lenses”. In: *Optics Express* 22.16 (Aug. 2014), pp. 19293–19301. DOI: 10.1364/OE.22.019293.
- [Fav05] FAVARO, P. and SOATTO, S.: “A geometric approach to shape from defocus”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 27.3 (Mar. 2005), pp. 406–417. DOI: 10.1109/TPAMI.2005.43.
- [Gmi93] GMITRO, Arthur F. and AZIZ, David: “Confocal microscopy through a fiber-optic imaging bundle”. In: *Opt. Lett.* 18.8 (Apr. 1993), pp. 565–567. DOI: 10.1364/OL.18.000565.
- [Goo10] GOODMAN, Jonathan and WEARE, Jonathan: “Ensemble samplers with affine invariance”. In: *Communications in applied mathematics and computational science* 5.1 (2010), pp. 65–80.

- [Hag07] HAGEN, Guy M.; CAARLS, Wouter; THOMAS, Martin; HILL, Andrew; LIDKE, Keith A.; RIEGER, Bernd; FRITSCH, Cornelia; GEEST, Bert van; JOVIN, Thomas M. and ARNDT-JOVIN, Donna J.: “Biological applications of an LCoS-based programmable array microscope (PAM)”. In: *Proc. SPIE* 6441 (2007). DOI: 10.1117/12.710995.
- [Han01] HANLEY, Q. S.; VERVEER, P. J.; GEMKOW, M. J.; ARNDT-JOVIN, D. and JOVIN, T. M.: “An optical sectioning programmable array microscope implemented with a digital micromirror device”. In: *Journal of Microscopy* 196.3 (2001), pp. 317–331. DOI: 10.1046/j.1365-2818.1999.00602.x.
- [Haz18] HAZIRBAS, C.; SOYER, S. G.; STAAB, M. C.; LEAL-TAIXÉ, L. and CREMERS, D.: “Deep Depth From Focus”. In: *Asian Conference on Computer Vision (ACCV)*. Dec. 2018.
- [Hil12a] HILLENBRAND, M.; GREWE, A.; BICHRA, M.; MITSCHUNAS, B.; KIRNER, R.; WEIß, R. and SINZINGER, S.: “Chromatic information coding in optical systems for hyperspectral imaging and chromatic confocal sensing”. In: *Proc.SPIE* 8550 (2012), p. 10. DOI: 10.1117/12.981147.
- [Hil12b] HILLENBRAND, Matthias; MITSCHUNAS, Beate; WENZEL, Christian; GREWE, Adrian; MA, Xuan; FEßER, Patrick; BICHRA, Mohamed and SINZINGER, Stefan: “Hybrid hyperchromats for chromatic confocal sensor systems”. In: *Advanced Optical Technologies* 1.3 (2012), pp. 187–194.
- [Hil13a] HILLENBRAND, Matthias; GREWE, Adrian; BICHRA, Mohamed; KLEINDIENST, Roman; LORENZ, Lucia; KIRNER, Raoul; WEIß, Robert and SINZINGER, Stefan: “Parallelized chromatic confocal sensor systems”. In: *Proc.SPIE* 8788 (2013), p. 10. DOI: 10.1117/12.2020334.
- [Hil13b] HILLENBRAND, Matthias; LORENZ, Lucia; KLEINDIENST, Roman; GREWE, Adrian and SINZINGER, Stefan: “Spectrally multiplexed chromatic confocal multipoint sensing”. In: *Opt. Lett.* 38.22 (Nov. 2013), pp. 4694–4697. DOI: 10.1364/OL.38.004694.

- [Hil15] HILLENBRAND, Matthias; WEISS, Robert; ENDRÖDY, Csaba; GREWE, Adrian; HOFFMANN, Martin and SINZINGER, Stefan: “Chromatic confocal matrix sensor with actuated pinhole arrays”. In: *Appl. Opt.* 54.15 (May 2015), pp. 4927–4936. DOI: 10.1364/AO.54.004927.
- [Hir16] HIRAI, Keita; IRIE, Daisuke and HORIUCHI, Takahiko: “Multi-primary image projector using programmable spectral light source”. In: *Journal of the Society for Information Display* 24.3 (2016), pp. 144–153. DOI: 10.1002/jsid.422.
- [Hoc97] HOCHREITER, S. and SCHMIDHUBER, J.: “Long Short-Term Memory”. In: *Neural Computation* 9.8 (Nov. 1997), pp. 1735–1780. DOI: 10.1162/neco.1997.9.8.1735.
- [Hua96] HUANG, S. H.; ZOU, X. Y.; HWANG, S. M.; WILLNER, A. E.; BAO, Z. and SMITH, D. A.: “Experimental demonstration of dynamic network equalization of three 2.5-Gb/s WDM channels over 1000 km using acoustooptic tunable filters”. In: *IEEE Photonics Technology Letters* 8.9 (Sept. 1996), pp. 1243–1245. DOI: 10.1109/68.531850.
- [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* 7.0 (2012).
- [Hut88] HUTLEY, M C and STEVENS, R F: “The use of a zone-plate monochromator as a displacement transducer”. In: *Journal of Physics E: Scientific Instruments* 21.11 (Nov. 1988), pp. 1037–1044. DOI: 10.1088/0022-3735/21/11/007.
- [Inu01] INUI, T.; KOMUKAI, T. and NAKAZAWA, M.: “Highly efficient tunable fiber Bragg grating filters using multilayer piezoelectric transducers”. In: *Optics Communications* 190.1–6 (2001), pp. 1–4. DOI: 10.1016/S0030-4018(01)00987-7.
- [Jeo16] JEONG, Hyeong-jun; YOO, Hongki and GWEON, DaeGab: “High-speed 3-D measurement with a large field of view based on direct-view confocal microscope with an electrically tunable lens”. In: *Opt. Express* 24.4 (Feb. 2016), pp. 3806–3816. DOI: 10.1364/OE.24.003806.

- [Jon93] JONES, G R and RUSSELL, P C: “Chromatic modulation based metrology”. In: *Pure and Applied Optics: Journal of the European Optical Society Part A 2.2* (Mar. 1993), pp. 87–110. DOI: 10.1088/0963-9659/2/2/003.
- [Kan18] KANG, Jian; LI, Xiang and LIU, Quan: “Hadamard transform-based calibration method for programmable optical filters based on digital micro-mirror device”. In: *Opt. Express* 26.15 (July 2018), pp. 19563–19573. DOI: 10.1364/OE.26.019563.
- [Kap01] KAPLAN, Ariel; FRIEDMAN, Nir and DAVIDSON, Nir: “Acousto-optic lens with very fast focus scanning”. In: *Opt. Lett.* 26.14 (July 2001), pp. 1078–1080. DOI: 10.1364/OL.26.001078.
- [Kha06] KHAN, Sajjad A. and RIZA, Nabeel A.: “Demonstration of a no-moving-parts axial scanning confocal microscope using liquid crystal optics”. In: *Optics Communications* 265.2 (2006), pp. 461–467. DOI: 10.1016/j.optcom.2006.03.068.
- [Kim06] KIM, Jeongmin; KANG, DongKyun and GWEON, DaeGab: “Spectrally encoded slit confocal microscopy”. In: *Optics Letters* 31.11 (June 2006), p. 1687. DOI: 10.1364/OL.31.001687.
- [Kim13] KIM, Taejoong; KIM, Sang Hoon; DO, DukHo; YOO, Hongki and GWEON, DaeGab: “Chromatic confocal microscopy with a novel wavelength detection method using transmittance”. In: *Opt. Express* 21.5 (Mar. 2013), pp. 6286–6294. DOI: 10.1364/OE.21.006286.
- [Kim15] KIM, Soocheol; HWANG, Jaehyun; HEO, Jung; RYU, Su-Ho; LEE, Donghak; KIM, Sang-Hoon; OH, Seung Jae and JOO, Chulmin: “Spectrally encoded slit confocal microscopy using a wavelength-swept laser”. In: *Journal of Biomedical Optics* 20.3 (Mar. 2015), p. 036016. DOI: 10.1117/1.JBO.20.3.036016.
- [Kin14] KING, Sharon V.; DOBLAS, Ana; PATWARY, Nurmohammed; SAAVEDRA, Genaro; MARTÍNEZ-CORRAL, Manuel and PREZA, Chrysanthe: “Implementation of PSF engineering in high-resolution 3D microscopy imaging with a LCoS (reflective) SLM”. In: *Proc. SPIE* 8949 (2014). DOI: 10.1117/12.2040723.

- [Kou14] KOUKOURAKIS, Nektarios; FINKELDEY, Markus; STÜRMER, Moritz; LEITHOLD, Christoph; GERHARDT, Nils C.; HOFMANN, Martin R.; WALLRABE, Ulrike; CZARSKE, Jürgen W. and FISCHER, Andreas: “Axial scanning in confocal microscopy employing adaptive lenses (CAL)”. In: *Optics Express* 22.5 (Mar. 2014), p. 6025. DOI: 10.1364/OE.22.006025.
- [Lee14] LEE, Dong-Ryoung; KIM, Young-Duk; GWEON, Dae-Gab and YOO, Hongki: “High speed 3D surface profile without axial scanning: dual-detection confocal reflectance microscopy”. In: *Measurement Science and Technology* 25.12 (Oct. 2014), p. 125403. DOI: 10.1088/0957-0233/25/12/125403.
- [LEE18] LEE, D.; GWEON, D.-G. and YOO, H.: “Multipoint scanning dual-detection confocal microscopy for fast 3D volumetric measurement”. In: *Journal of Microscopy* 270.2 (2018), pp. 200–209. DOI: 10.1111/jmi.12674.
- [Li17] LI, Jizhou; XUE, Feng and BLU, Thierry: “Fast and accurate three-dimensional point spread function computation for fluorescence microscopy”. In: *J. Opt. Soc. Am. A* 34.6 (June 2017), pp. 1029–1034. DOI: 10.1364/JOSAA.34.001029.
- [Lin98] LIN, Paul C.; SUN, Pang-Chen; ZHU, Lijun and FAINMAN, Yeshaiah: “Single-shot depth-section imaging through chromatic slit-scan confocal microscopy”. In: *Appl. Opt.* 37.28 (Oct. 1998), pp. 6764–6770. DOI: 10.1364/AO.37.006764.
- [Liu14] LIU, Lin; WANG, Erkang; ZHANG, Xiaoyang; LIANG, Wenxuan; LI, Xingde and XIE, Huikai: “MEMS-based 3D confocal scanning microendoscope using MEMS scanners for both lateral and axial scan”. In: *Sensors and Actuators A: Physical* 215 (2014), pp. 89–95. DOI: 10.1016/j.sna.2013.09.035.
- [Luo12] LUO, Ding; KUANG, Cuifang and LIU, Xu: “Fiber-based chromatic confocal microscope with Gaussian fitting method”. In: *Optics & Laser Technology* 44.4 (2012), pp. 788–793. DOI: 10.1016/j.optlastec.2011.10.027.

- [Luo16a] LUO, Ding: “Linear Model for Optical Measurement”. In: *Proceedings of the 2015 Joint Workshop of Fraunhofer IOSB and Institute for Anthropomatics, Vision and Fusion Laboratory*. Vol. 24. KIT Scientific Publishing. 2016, p. 17.
- [Luo16b] LUO, Ding; BAUER, Sebastian; TAPHANEL, Miro; LÄNGLE, Thomas; LEÓN, Fernando Puente and BEYERER, Jürgen: “Optical unmixing using programmable spectral source based on DMD”. In: *Next-Generation Spectroscopic Technologies IX*. Vol. 9855. International Society for Optics and Photonics. 2016, 98550P. DOI: 10.1117/12.2228303.
- [Luo16c] LUO, Ding; LÄNGLE, Thomas and BEYERER, Jürgen: “Compressive shape from focus based on a linear measurement model”. In: *Forum Bildverarbeitung 2016*. Karlsruhe, Germany, Dec. 2016, pp. 185–195.
- [Luo17a] LUO, Ding: “RNN-accelerated Experimental Design for Chromatic Confocal Measurement”. In: *Proceedings of the 2016 Joint Workshop of Fraunhofer IOSB and Institute for Anthropomatics, Vision and Fusion Laboratory*. Vol. 33. KIT Scientific Publishing. 2017, p. 17.
- [Luo17b] LUO, Ding; LÄNGLE, Thomas and BEYERER, Jürgen: “Compressive shape from focus based on a linear measurement model”. In: *tm-Technisches Messen* 84.7-8 (2017), pp. 452–459. DOI: 10.1515/teme-2017-0007.
- [Luo17c] LUO, Ding; TAPHANEL, Miro; LÄNGLE, Thomas and BEYERER, Jürgen: “Programmable light source based on an echellogram of a supercontinuum laser”. In: *Appl. Opt.* 56.8 (Mar. 2017), pp. 2359–2367. DOI: 10.1364/AO.56.002359.
- [Luo18] LUO, Ding: “Adaptive Measurement Method for Area Chromatic Confocal Microscopy”. In: *Proceedings of the 2017 Joint Workshop of Fraunhofer IOSB and Institute for Anthropomatics, Vision and Fusion Laboratory*. Vol. 34. KIT Scientific Publishing. 2018, p. 31.

- [Luo20] LUO, Ding; TAPHANEL, Miro; CLAUS, Daniel; BOETTCHER, Tobias; OSTEN, Wolfgang; LÄNGLE, Thomas and BEYERER, Jürgen: “Area scanning method for 3D surface profilometry based on an adaptive confocal microscope”. In: *Optics and Lasers in Engineering* 124 (2020), p. 105819. DOI: 10.1016/j.optlaseng.2019.105819.
- [Mah13] MAHMOOD, M.T.: “Shape From Focus by Total Variation”. In: *IVMSP Workshop, 2013 IEEE 11th*. June 2013, pp. 1–4. DOI: 10.1109/IVMSPW.2013.6611940.
- [Mar15] MARTÍN ABADI et al.: *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. 2015.
- [Mar76] MARR, D. and POGGIO, T.: *Cooperative Computation of Stereo Disparity*. Tech. rep. Cambridge, MA, USA, 1976.
- [Men97] MENDLOVIC, David; GARCIA, Javier; ZALEVSKY, Zeev; MAROM, Emanuel; MAS, David; FERREIRA, Carlos and LOHMANN, Adolf W.: “Wavelength-multiplexing system for single-mode image transmission”. In: *Applied Optics* 36.32 (Nov. 1997), p. 8474. DOI: 10.1364/AO.36.008474.
- [Mer08] MERMILLOD-BLONDIN, Alexandre; MCLEOD, Euan and ARNOLD, Craig B.: “High-speed varifocal imaging with a tunable acoustic gradient index of refraction lens”. In: *Optics Letters* 33.18 (Sept. 2008), p. 2146. DOI: 10.1364/OL.33.002146.
- [Min09] MINHAS, Rashid; MOHAMMED, Abdul A.; WU, Q. M. Jonathan and SID-AHMED, Maher A.: “3D Shape from Focus and Depth Map Computation Using Steerable Filters”. In: *Image Analysis and Recognition: 6th International Conference, ICIAR 2009, Halifax, Canada, July 6-8, 2009. Proceedings*. Ed. by KAMEL, Mohamed and CAMPILHO, Aurélio. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 573–583. DOI: 10.1007/978-3-642-02611-9_57.
- [Min61] MINSKY, Marvin: “Microscopy apparatus”. US Patent 3,013,467. Dec. 1961.

- [Mol84] MOLESINI, G.; PEDRINI, G.; POGGI, P. and QUERCIOLO, F.: “Focus-wavelength encoded optical profilometer”. In: *Optics Communications* 49.4 (1984), pp. 229–233. DOI: 10.1016/0030-4018(84)90179-2.
- [Nay89] NAYAR, Shree K: Shape from focus. Tech. rep. CARNEGIE-MELLON UNIV PITTSBURGH PA ROBOTICS INST, 1989.
- [Nay94] NAYAR, S.K. and NAKAGAWA, Y.: “Shape from focus”. In: *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 16.8 (Aug. 1994), pp. 824–831. DOI: 10.1109/34.308479.
- [Pap06] PAPASTATHOPOULOS, Evangelos; KÖRNER, Klaus and OSTEN, Wolfgang: “Chromatic confocal spectral interferometry”. In: *Appl. Opt.* 45.32 (Nov. 2006), pp. 8244–8252. DOI: 10.1364/AO.45.008244.
- [Par98] PARKER, M. C.; COHEN, A. D. and MEARS, R. J.: “Dynamic digital holographic wavelength filtering”. In: *Journal of Lightwave Technology* 16.7 (July 1998), pp. 1259–1270. DOI: 10.1109/50.701404.
- [Pat91] PATEL, J. S. and MAEDA, M. W.: “Tunable polarization diversity liquid-crystal wavelength filter”. In: *IEEE Photonics Technology Letters* 3.8 (Aug. 1991), pp. 739–740. DOI: 10.1109/68.84470.
- [Per13] PERTUZ, Said; PUIG, Domenec and GARCIA, Miguel Angel: “Analysis of Focus Measure Operators for Shape-from-focus”. In: *Pattern Recogn.* 46.5 (May 2013), pp. 1415–1432. DOI: 10.1016/j.patcog.2012.11.011.
- [Pit03] PITRIS, C.; BOUMA, B. E.; SHISKOV, M. and TEARNEY, G. J.: “A GRISM-based probe for spectrally encoded confocal microscopy”. In: *Optics Express* 11.2 (Jan. 2003), pp. 120–124. DOI: 10.1364/OE.11.000120.
- [Poh08] POHER, Vincent; KENNEDY, Gordon T.; MANNING, Hugh B.; OWEN, Dylan M.; ZHANG, Haoxiang X.; GU, Erdan; DAWSON, Martin D.; FRENCH, Paul M. W. and NEIL, Mark A. A.: “Improved sectioning in a slit scanning confocal microscope”. In: *Opt. Lett.* 33.16 (Aug. 2008), pp. 1813–1815. DOI: 10.1364/OL.33.001813.

- [Ray14] RAYER, Mathieu and MANSFIELD, Daniel: “Chromatic confocal microscope using hybrid aspheric diffractive lenses”. In: *Proc.SPIE* 9130 (2014), p. 7. DOI: 10.1117/12.2048691.
- [Riz03] RIZA, Nabeel A. and MUGHAL, M. Junaid: “Broadband optical equalizer using fault-tolerant digital micromirrors”. In: *Opt. Express* 11.13 (June 2003), pp. 1559–1565. DOI: 10.1364/OE.11.001559.
- [Riz99] RIZA, Nabeel A. and SUMRIDDETKHAJORN, Sarun: “Digitally controlled fault-tolerant multiwavelength programmable fiber-optic attenuator using a two-dimensional digital micromirror device”. In: *Opt. Lett.* 24.5 (Mar. 1999), pp. 282–284. DOI: 10.1364/OL.24.000282.
- [Rya03] RYAN, Kenneth J: “Estimating Expected Information Gains for Experimental Designs With Application to the Random Fatigue-Limit Model”. In: *Journal of Computational and Graphical Statistics* 12.3 (2003), pp. 585–603. DOI: 10.1198/1061860032012.
- [Sab99] SABHARWAL, Yashvinder S.; ROUSE, Andrew R.; DONALDSON, La-Tanya; HOPKINS, Mark F. and GMITRO, Arthur F.: “Slit-scanning confocal microendoscope for high-resolution in vivo imaging”. In: *Applied Optics* 38.34 (Dec. 1999), p. 7133. DOI: 10.1364/AO.38.007133.
- [San97] SANTOS, A.; ORTIZ DE SOLORZANO, C.; VAQUERO, J. J.; PEÑA, J. M.; MALPICA, N. and DEL POZO, F.: “Evaluation of autofocus functions in molecular cytogenetic analysis”. In: *Journal of microscopy* 188.3 (Dec. 1997), pp. 264–272. DOI: 10.1046/j.1365-2818.1997.2630819.x.
- [Sap02] SAPRIEL, J.; CHARISSOUX, D.; VOLOSHINOV, V. and MOLCHANOV, V.: “Tunable acoustooptic filters and equalizers for WDM applications”. In: *Journal of Lightwave Technology* 20.5 (May 2002), pp. 892–899. DOI: 10.1109/JLT.2002.1007946.
- [Sch11] SCHAFFER, Martin; GROSSE, Marcus; HARENDT, Bastian and KOWARSCHIK, Richard: “High-speed three-dimensional shape measurements of objects with laser speckles and acousto-optical deflection”. In: *Opt. Lett.* 36.16 (Aug. 2011), pp. 3097–3099. DOI: 10.1364/OL.36.003097.

- [She88] SHEPPARD, Colin and MAO, X.Q.: “Confocal Microscopes with Slit Apertures”. In: *Journal of Modern Optics* 35 (Mar. 1988), pp. 1169–1185. DOI: 10.1080/09500348814551251.
- [Shi04] SHI, Kebin; LI, Peng; YIN, Shizhuo and LIU, Zhiwen: “Chromatic confocal microscopy using supercontinuum light”. In: *Opt. Express* 12.10 (May 2004), pp. 2096–2101. DOI: 10.1364/OPEX.12.002096.
- [Shi92] SHIBAGUCHI, Takashi and FUNATO, Hiroyoshi: “Lead-Lanthanum Zirconate-Titanate (PLZT) Electrooptic Variable Focal-Length Lens with Stripe Electrodes”. In: *Japanese Journal of Applied Physics* 31.Part 1, No. 9B (Sept. 1992), pp. 3196–3200. DOI: 10.1143/jjap.31.3196.
- [Sie95] SIEGELMANN, Hava T.: “Computation Beyond the Turing Limit”. In: *Science* 268.5210 (1995), pp. 545–548. DOI: 10.1126/science.268.5210.545.
- [Sin15] SINHARROY, Indranil; XY124; HOLLOWAY, Catherine; NG110 and NUMMELA, Ville: PyZDDE: v1.0-alpha. Mar. 2015. DOI: 10.5281/zenodo.15763.
- [Sos11] SOSKIND, Y.G.: *Field Guide to Diffractive Optics*. Field Guide Series no. 1. SPIE, 2011.
- [Sub95] SUBBARAO, Murali and CHOI, Tae: “Accurate Recovery of Three-Dimensional Shape from Image Focus”. In: *IEEE Trans. Pattern Anal. Mach. Intell.* 17.3 (Mar. 1995), pp. 266–274. DOI: 10.1109/34.368191.
- [Szu18] SZULZYCKI, Krzysztof; SAVARYN, Viktoriya and GRULKOWSKI, Ireneusz: “Rapid acousto-optic focus tuning for improvement of imaging performance in confocal microscopy [Invited]”. In: *Applied Optics* 57.10 (Apr. 2018), pp. C14–C18. DOI: 10.1364/AO.57.000C14.
- [Tap13] TAPHANEL, Miro; HOVESTREYDT, Bastiaan and BEYERER, Jürgen: “Speed-up chromatic sensors by optimized optical filters”. In: *Proc.SPIE* 8788 (2013), p. 10. DOI: 10.1117/12.2020387.

- [Tea98] TEARNEY, G. J.; WEBB, R. H. and BOUMA, B. E.: “Spectrally encoded confocal microscopy”. In: *Optics Letters* 23.15 (Aug. 1998), pp. 1152–1154. DOI: 10.1364/OL.23.001152.
- [The09] THELEN, A.; FREY, S.; HIRSCH, S. and HERING, P.: “Improvements in Shape-From-Focus for Holographic Reconstructions With Regard to Focus Operators, Neighborhood-Size, and Height Value Interpolation”. In: *IEEE Transactions on Image Processing* 18.1 (Jan. 2009), pp. 151–157. DOI: 10.1109/TIP.2008.2007049.
- [Tiz94a] TIZIANI, H. J. and UHDE, H.-M.: “Three-dimensional image sensing by chromatic confocal microscopy”. In: *Applied Optics* 33.10 (Apr. 1994), pp. 1838–1843. DOI: 10.1364/AO.33.001838.
- [Tiz94b] TIZIANI, Hans J. and UHDE, Hans-Martin: “Three-dimensional analysis by a microlens-array confocal arrangement”. In: *Appl. Opt.* 33.4 (Feb. 1994), pp. 567–572. DOI: 10.1364/AO.33.000567.
- [Tiz96] TIZIANI, H. J.; ACHI, R. and KRÄMER, R. N.: “Chromatic confocal microscopy with microlenses”. In: *Journal of Modern Optics* 43.1 (1996), pp. 155–163. DOI: 10.1080/09500349608232730.
- [van07] VAN DEN BOS, Adriaan: *Parameter Estimation for Scientists and Engineers*. New York, NY, USA: Wiley-Interscience, 2007.
- [Vav10] VAVASIS, S.: “On the Complexity of Nonnegative Matrix Factorization”. In: *SIAM Journal on Optimization* 20.3 (2010), pp. 1364–1377. DOI: 10.1137/070709967.
- [Vu12] VU, C.T.; PHAN, T.D. and CHANDLER, D.M.: “ S_3 : A Spectral and Spatial Measure of Local Perceived Sharpness in Natural Images”. In: *Image Processing, IEEE Transactions on* 21.3 (Mar. 2012), pp. 934–945. DOI: 10.1109/TIP.2011.2169974.
- [Wil84] WILSON, Tony and SHEPPARD, Colin: *Theory and practice of scanning optical microscopy*. London ; Orlando: Academic Press, 1984.
- [Woo12] WOOD, Tobias C and ELSON, Daniel S: “A tunable supercontinuum laser using a digital micromirror device”. In: *Measurement Science and Technology* 23.10 (2012), p. 105204.

- [Xia88] XIAO, G. Q.; CORLE, T. R. and KINO, G. S.: “Real-time confocal scanning optical microscope”. In: *Applied Physics Letters* 53.8 (Aug. 1988), pp. 716–718. DOI: 10.1063/1.99814.
- [Zha00] ZHANG, Zhengyou: “A flexible new technique for camera calibration”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22.11 (Nov. 2000), pp. 1330–1334. DOI: 10.1109/34.888718.
- [Zha99] ZHANG, Ruo; TSAI, P.-S.; CRYER, J.E. and SHAH, M.: “Shape-from-shading: a survey”. In: *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 21.8 (Aug. 1999), pp. 690–706. DOI: 10.1109/34.784284.

Publications

- [Luo12] LUO, Ding; KUANG, Cuifang and LIU, Xu: “Fiber-based chromatic confocal microscope with Gaussian fitting method”. In: *Optics & Laser Technology* 44.4 (2012), pp. 788–793. DOI: 10.1016/j.optlastec.2011.10.027.
- [Bau16] BAUER, Sebastian; KRIPPNER, Wolfgang; LUO, Ding; TAPHANEL, Miro; ABDO, Mohammad; BADILITA, Vlad; LÄNGLE, Thomas; BEYERER, Jürgen; KORVINK, Jan and LEÓN, Fernando Puente: “Ortsaufgelöste optische Bestimmung von Materialanteilen in Mischungen”. In: *XXX. Messtechnisches Symposium*. Walter de Gruyter GmbH & Co KG. 2016, p. 69.
- [Luo16a] LUO, Ding: “Linear Model for Optical Measurement”. In: *Proceedings of the 2015 Joint Workshop of Fraunhofer IOSB and Institute for Anthropomatics, Vision and Fusion Laboratory*. Vol. 24. KIT Scientific Publishing. 2016, p. 17.
- [Luo16b] LUO, Ding; BAUER, Sebastian; TAPHANEL, Miro; LÄNGLE, Thomas; LEÓN, Fernando Puente and BEYERER, Jürgen: “Optical unmixing using programmable spectral source based on DMD”. In: *Next-Generation Spectroscopic Technologies IX*. Vol. 9855. International Society for Optics and Photonics. 2016, 98550P. DOI: 10.1117/12.2228303.
- [Luo16c] LUO, Ding; LÄNGLE, Thomas and BEYERER, Jürgen: “Compressive shape from focus based on a linear measurement model”. In: *Forum Bildverarbeitung 2016*. Karlsruhe, Germany, Dec. 2016, pp. 185–195.

- [Qu16] QU, Chengchao; LUO, Ding; MONARI, Eduardo; SCHUCHERT, Tobias and BEYERER, Jürgen: “Capturing ground truth super-resolution data”. In: *2016 IEEE International Conference on Image Processing (ICIP)*. IEEE. 2016, pp. 2812–2816.
- [Luo17a] LUO, Ding: “RNN-accelerated Experimental Design for Chromatic Confocal Measurement”. In: *Proceedings of the 2016 Joint Workshop of Fraunhofer IOSB and Institute for Anthropomatics, Vision and Fusion Laboratory*. Vol. 33. KIT Scientific Publishing. 2017, p. 17.
- [Luo17b] LUO, Ding; LÄNGLE, Thomas and BEYERER, Jürgen: “Compressive shape from focus based on a linear measurement model”. In: *tm-Technisches Messen* 84.7-8 (2017), pp. 452–459. DOI: 10.1515/teme-2017-0007.
- [Luo17c] LUO, Ding; TAPHANEL, Miro; LÄNGLE, Thomas and BEYERER, Jürgen: “Programmable light source based on an echellogram of a supercontinuum laser”. In: *Appl. Opt.* 56.8 (Mar. 2017), pp. 2359–2367. DOI: 10.1364/AO.56.002359.
- [Luo18] LUO, Ding: “Adaptive Measurement Method for Area Chromatic Confocal Microscopy”. In: *Proceedings of the 2017 Joint Workshop of Fraunhofer IOSB and Institute for Anthropomatics, Vision and Fusion Laboratory*. Vol. 34. KIT Scientific Publishing. 2018, p. 31.
- [Luo20] LUO, Ding; TAPHANEL, Miro; CLAUS, Daniel; BOETTCHER, Tobias; OSTEN, Wolfgang; LÄNGLE, Thomas and BEYERER, Jürgen: “Area scanning method for 3D surface profilometry based on an adaptive confocal microscope”. In: *Optics and Lasers in Engineering* 124 (2020), p. 105819. DOI: 10.1016/j.optlaseng.2019.105819.

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

- Band 1 **JONATHAN BALZER**
Regularisierung des Deflektometrieproblems Grundlagen
und Anwendung. 2008
ISBN 978-3-86644-230-6
- Band 2 **IOANA GHEȚA**
Fusion multivariater Bildserien am Beispiel eines Kamera-Arrays. 2011
ISBN 978-3-86644-684-7
- Band 3 **STEFAN BRUNO WERLING**
Deflektometrie zur automatischen Sichtprüfung
und Rekonstruktion spiegelnder Oberflächen. 2011
ISBN 978-3-86644-687-8
- Band 4 **JAN WASSENBERG**
Efficient Algorithms for Large-Scale Image Analysis. 2012
ISBN 978-3-86644-786-8
- Band 5 **MARTIN GRAFMÜLLER**
Verfahrensfortschritte in der robusten Echtzeiterkennung
von Schriftzeichen. 2013
ISBN 978-3-86644-979-4
- Band 6 **JÜRGEN BRAUER**
Human Pose Estimation with Implicit Shape Models. 2014
ISBN 978-3-7315-0184-8
- Band 7 **MARKUS MÜLLER**
Szeneninterpretation unter Verwendung multimodaler Sensorik
und Salienzmaßen. 2017
ISBN 978-3-7315-0240-1
- Band 8 **ROBIN GRUNA**
Beleuchtungsverfahren zur problemspezifischen Bildgewinnung
für die automatische Sichtprüfung. 2015
ISBN 978-3-7315-0313-2
- Band 9 **THOMAS STEPHAN**
Beitrag zur Unterwasserbildrestauration. 2017
ISBN 978-3-7315-0579-2

- Band 10 **JAN-PHILIP JARVIS**
A Contribution to Active Infrared Laser Spectroscopy
for Remote Substance Detection. 2017
ISBN 978-3-7315-0725-3
- Band 11 **MIRO TAPHANEL**
Chromatisch konfokale Triangulation – Hochgeschwindigkeits 3D-Sensorik
auf Basis der Wellenlängenschätzung mit optimierten Filtern. 2018
ISBN 978-3-7315-0646-1
- Band 12 **SEBASTIAN HÖFER**
Untersuchung diffus spiegelnder Oberflächen mittels
Infrarotdeflektometrie. 2017
ISBN 978-3-7315-0711-6
- Band 13 **MATTHIAS RICHTER**
Über lernende optische Inspektion am Beispiel der
Schüttgutsortierung. 2018
ISBN 978-3-7315-0842-7
- Band 14 **MATHIAS ZIEBARTH**
Wahrnehmungsgrenzen kleiner Verformungen auf
spiegelnden Oberflächen. 2019
ISBN 978-3-7315-0890-8
- Band 15 **JOHANNES MEYER**
Light Field Methods for the Visual Inspection of Transparent Objects. 2019
ISBN 978-3-7315-0912-7
- Band 16 **MASOUD ROSCHANI**
Probabilistische Planungsverfahren für die
deflektometrische Oberflächeninspektion. 2019
ISBN 978-3-7315-0907-3
- Band 17 **MAHSA MOHAMMADIKAJI**
Simulation-based Planning of Machine Vision Inspection Systems
with an Application to Laser Triangulation. 2020
ISBN 978-3-7315-0989-9
- Band 18 **DING LUO**
High-speed surface profilometry based on an adaptive microscope
with axial chromatic encoding. 2021
ISBN 978-3-7315-1061-1

For the quality assurance of a technical part, the 3D geometric profile of the working surface is often one of the most important aspects, which directly affects the functionality of the part in a fundamental way. Over the past decades, optical 3D surface profilometry has gained an increasing amount of attention for inspection tasks in both academic and industrial environments, due to its capability of non-contact measurement and high resolution.

An adaptive microscope with axial chromatic encoding is designed and developed, namely the AdaScope. With a holistic design approach, the AdaScope consists of two major components. Firstly, the programmable light source is based on a supercontinuum laser, whose echellogram is spatially filtered by a digital micromirror device (DMD). By sending different patterns to the DMD, arbitrary spectra can be generated for the output light. Secondly, the programmable array microscope is constructed based on a second DMD, which serves as a programmable array of secondary light source. A chromatic objective is utilized so that the necessity of axial mechanical scanning is avoided. The combination of both components grants the AdaScope the ability to confocally address any locations within the measurement volume, which provides the hardware foundation for a cascade measurement strategy to be developed, dramatically accelerating the speed of 3D confocal microscopy.

ISSN 1866-5934

ISBN 978-3-7315-1061-1

Gedruckt auf FSC-zertifiziertem Papier

