APPLICATIONS OF DIGITAL HOLOGRAPHY IN DIRECT PHASE RETRIEVAL USING TRANSPORT OF INTENSITY AND IN 3D SURFACE FEATURE EXTRACTION

Thesis

Submitted to

The School of Engineering of the

UNIVERSITY OF DAYTON

In Partial Fulfillment of the Requirements for

The Degree of

Master of Science in Electro-Optics

By

Haowen Zhou

UNIVERSITY OF DAYTON

Dayton, Ohio

May 2021

APPLICATIONS OF DIGITAL HOLOGRAPHY IN DIRECT PHASE RETRIEVAL USING TRANSPORT OF INTENSITY AND IN 3D SURFACE FEATURE EXTRACTION

Name: Zhou, Haowen

APPROVED BY:

Partha P. Banerjee
Advisory Committee Chairman
Professor
Department of Electro-Optics and Photonics

Andrew Sarangan
Committee Member
Professor and Chair
Department of Electro-Optics and Photonics

Chenglong Zhao
Committee Member
Associate Professor
Department of Physics

Robert J. Wilkens, Ph.D., P.E.
Associate Dean for Research and Innovation
Professor
School of Engineering

Eddy M. Rojas, Ph.D., M.A., P.E.
Dean, School of Engineering

ABSTRACT

APPLICATIONS OF DIGITAL HOLOGRAPHY IN DIRECT PHASE RETRIEVAL USING TRANSPORT OF INTENSITY AND IN 3D SURFACE FEATURE EXTRACTION

Name: Zhou, Haowen
University of Dayton

Advisor: Dr. Partha P. Banerjee

Phase information plays an important role in three-dimensional (3D) imaging. Digital holography (DH) has been widely used in 3D imaging to reconstruct the phase distribution from optical fields. The transport of intensity equation (TIE) is a way to retrieve true or unwrapped phase from phase objects with multiple captures of intensity images along the light propagation direction. In this thesis, solving TIE with the help of the transport of phase equation (TPE) is first proposed. Combining TIE with TPE can further improve the accuracy of the Fourier transform based TIE solution with no extra captures of intensity images and little sacrifice of computational efforts. Next, the combination of DH and TIE is investigated to retrieve phase without any phase unwrapping and multiple captures of intensities. Finally, correlation of holograms is proposed to extract 3D surface features directly from 2D holograms. As an illustration, an application to face recognition is presented.

Dedicated to my parents

For their unconditional love all the time

ACKNOWLEDGMENTS

My special thanks are to Department of Electro-Optics and Photonics, and School of Engineering at University of Dayton who supported me financially for my graduate study. I would like to sincerely thank my advisor Dr. Partha Banerjee for his devotion, inspiring guidance, detailed instructions, continuous encouragement, and full support throughout my graduate work and study. I would also like to thank Dr. Elena Stoykova from the Bulgarian Academy of Sciences and Dr. Liangcai Cao from Tsinghua University for their support, instructions, and fruitful discussions on this work. I truly appreciate my committee members Dr. Andrew Sarangan and Dr. Chenglong Zhao for their constructive suggestions on this work. I would like to express my thanks to all my course lecturers. Without them, I can never build solid knowledge foundation in optics and photonics.

I sincerely thank our group members, Dr. Ujitha Abeywickrema, Dr. Behzad Bordbar, Dr. Mallik Hussain, Yunyang Li, Haiyun Guo, Saleha Qissi for fruitful discussions on the work. I appreciate to have lots of friends and classmates in the Department of Electro-Optics and Photonics. Thanks for working together going through the COVID-19 situation. Among all friends and classmates, I want to extend my special thanks to my officer team of the SPIE/OSA student chapter, who gave me strong support in organizing academic events and social activities during this special period.

Last but not least, I especially thank my parents for giving me unconditional love and strong support all the time, and my cousin Kai Ye for giving me helpful suggestions on career plans and life plans.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

1D	One dimensional
2D	Two dimensional
3D	Three dimensional
BS	Beam splitter
CCD	Charged coupled device
CDI	Coherent diffraction imaging
CGH	Computer-generated hologram
CNN	Convolutional neural network
DH	Digital holography
DIC	Differential interference contrast
DR	Discrimination ratio
DWDH	Dual-wavelength digital holography
FFT	Fast Fourier transform
FPM	Fourier ptychography microscopy
GS	Gerchberg-Saxton
IFT	Inverse Fourier transform
LIGO	Laser interferometer gravitational-wave observatory
MAPE	Mean average pixel Error
MSE	Mean squared error

MWDH	Multi-wavelength digital holography
NME	Normalized mean error
PC	Personal computer
PCA	Principle component analysis
PCE	Peak-to-correlation plane energy
PNR	Peak-to-noise ratio
PSDH	Phase-shifting digital holography
PUMA	Phase unwrapping max-flow algorithm
RHS	Right-handed side
RMSE	Root mean squared error
SWDH	Single-wavelength digital holography
TIE	Transport of intensity equation
TPE	Transport of phase equation

LIST OF NOTATIONS

A_i	Constant amplitude of the i^{th} point source
A_o	Constant amplitude of the object field
A_r	Constant amplitude of the reference field
a	Amplitude of the envelope of the optical field
d	Propagation distance
E	Correlation plane energy
$\bar{E}$	Averaged correlation plane energy
E_e	Envelope of the of the optical field
$E_{od}(x,y)$	Diffracted object field at the plane $z = d$
$E_{o0}(x,y)$	Diffracted object field at the plane $z = 0$
$E_{o0}{}'(x,y)$	Reconstructed object field
E_p	Optical field phasor
$\mathcal{F}$	Fourier transform
$\mathcal{F}^{-1}$	Inverse Fourier transform
$H_{pi}(k_x,k_y)$	Angular spectrum transfer function
$h(x,y)$	Hologram function
$h_p(x,y;z)$	Impulse response of propagation
$h_{\phi_r}(x,y;d)$	Hologram function with phase shift of ϕ_r

I	Intensity of the optical field
$\bar{I}$	Approximated intensity at the focused reconstruction plane
$Im(\cdot)$	Imaginary part
$I_{1,0,-1}$	Intensity at the planes $z = -\Delta z, 0, \Delta z$
k_0	Wavenumber
$k_{0x,0y}$	Spatial frequency introduced by the tilt angle of the reference field
$k_{x,y}$	Transverse spatial frequencies in x, y
L	Length of the object plane
L_0	Length of feature points assembly in $x - y$ plane
M	Pixel number along x-direction
N	Pixel number along y-direction
R	Radius of curvature of the lens
$Re(\cdot)$	Real part
$\vec{r}_\perp$	Transverse spatial vector
P_{Auto}	Normalized autocorrelation peak value
P_{Cross}	Normalized cross-correlation peak value
$P_i(x_i, y_i, z_i)$	i^{th} point in object plane
Q_1	First quartile value of the data set
Q_2	Median value or second quartile value of the data set
Q_3	Third quartile value of the data set

W	Width of the object plane
W_0	Width of feature points assembly in $x-y$ plane
x_i	x axis location of i^{th} point source
y_i	y axis location of i^{th} point source
$\Gamma(x,y;z)$	Optical field at the plane z
Δx	Pixel pitch in x-direction
Δy	Pixel pitch in y-direction
Δz	Distance between two adjacent reconstruction planes
$\Delta\lambda$	Wavelength difference
$\theta_{x,y}$	Tilt angle of reference field along x- or y-axis
Λ	Synthetic wavelength
λ	Wavelength
σ	Width of the Gaussian phase
$\varphi(x,y)$	Phase of hologram
φ_λ	Measured phase with single wavelength digital holography
φ_Λ	Measured phase with dual wavelength digital holography
ϕ	Phase of the optical field
$\tilde{\phi}$	Phase of the optical field phasor
$\phi_o(x,y)$	Object phase distribution
$\phi_o{}'(x,y)$	Final retrieved object phase

ϕ_{om}	Designed maximum phase value of Gaussian phase
ϕ'_{om}	Retrieved maximum phase value of Gaussian phase
$\phi'_{0wrap}(x,y)$	Retrieved wrapped object phase
ϕ_r	Constant phase of the reference field

CHAPTER I

INTRODUCTION

1.1 Phase Objects and Phase Retrieval

Phase objects are common in our daily life, viz. window glass, window glass with raindrops, a glass of water, human cells, etc. When light transmits through these objects, the phase of the light is delayed by different amounts, while the amplitude of the light remains relatively unchanged during transmission. Similarly, when light is reflected off an opaque object, its phase can be spatially modulated depending on its depth or height profile. Phase retrieval has many applications, viz., in the biosciences for 3D imaging of phase objects [1-3], inspection of products in industry [4-5], in security for fingerprint and face recognition [6-7], to name a few.

In 1942, Zernike [8] invented the phase contrast microscope, which gave the first visualization of phase objects. In 1952, Normarski [9] invented the differential interference contrast (DIC) microscopy by employing the principle of polarized interference and shear interferometry. These methods have been well-developed and can give the qualitative phase profile.

To get a better understanding or illustration of phase distributions of phase objects, various methods for quantitative phase imaging has emerged over several decades. Based on the principle of interference, interferometry can be used to obtain phase information. The Laser Interferometer Gravitational-wave Observatory (LIGO) is a well-known example, which employs a giant Michelson interferometer [10]. Another quantitative phase imaging technique based on interference is holography, and more recently, digital holography (DH). The concept of holography was proposed by Gabor in 1948 [11]. With the advancement of computers and digital detectors, in 1967, Goodman and Lawrence introduced digital holograms, and phases were retrieved numerically [12]. More details about DH will be introduced in Section 1.3 and Section 3.1.

In addition to interferometric quantitative phase imaging techniques, there are also various non-interferometric methods. Iterative methods employ the intensity information at two different planes along the longitudinal (light propagation) axis. Then, phases are iteratively solved back and forth between two planes, with an error reduction algorithm [13]. In 1972, Gerchberg and Saxton proposed an efficient iterative phase retrieval algorithm [14], which is appropriately named the Gerchberg-Saxton (GS) algorithm. The GS algorithm has been widely used in coherent diffraction imaging (CDI) [15,16], adaptive optics [17,18], and computer-generated holography [19]. Ptychography is another iterative phase retrieval method proposed by Rodenburg and Farulker in 2004 [20,21]. In ptychography, the object is scanned by an optical field of finite size, and the diffraction patterns are recorded. In 2013, Zheng *et al.* [22] extended the concept of ptychography to the spatial frequency domain, which has now come to be called Fourier ptychography microscopy (FPM). Instead of illuminating the object with an aperture-constrained optical field, the object is illuminated by "plane waves" from different angles in high-resolution and for a wide field of view, thereby determining its angular plane wave spectrum.

In addition to iterative algorithms, another non-interferometric method is popular in phase retrieval, one that uses the transport of intensity equation (TIE). The equation is derived from the Helmholtz equation, which relates the intensity and the phase of the optical fields at each of the transverse plane. From the recorded intensity information at different distances of propagation, one can retrieve the phase of the optical field. More on TIE is discussed below and more extensively in Chapters II and III.

1.2 Transport of Intensity Equation

In the past four decades, transport of intensity has been widely studied and developed by using a sequence of intensity images along z-axis to solve for the true or unwrapped phase distribution [23]. TIE is derived by Teague from the Helmholtz equation for the complex optical field [24,25]. Independently, the coupled equations between amplitude and phase leading to the transport of

intensity and transport of phase equations (TIE, TPE) in a nonlinear Kerr medium was derived by Banerjee *et al*. [26] and applied to understanding self-focusing and defocusing of beams. The corresponding equations for propagation in a linear medium can be found in Banerjee and Poon [27]. These equations were an extension of the work by Akhmanov *et al*. [28] which now included the contribution from diffraction. The uniqueness of the solution of TIE was first proved by Gureyev *et al*. [29]. A variety of methods to solve the TIE have been proposed, viz., multigrid method [30], Zernike polynomial expansions [31], iterative methods [32,33], fast Fourier transform (FFT) based methods [33,34], discrete cosine transform method [35], recursive methods [36], finite difference methods [37,38], etc. Due to the advantages of simplicity and great time efficiency, the FFT-based method has been the most popular one, which is also adopted in our work. In the past two decades, TIE has been applied to objects illuminated with a partially coherent source [29,39,40], to biomedical imaging [41,42], to X-ray imaging [43], etc.

Avoiding a moving optical system required for multiple intensity registrations along the direction of propagation of the optical field to implement TIE can be of great convenience since it avoids registration issues stemming from the misalignment of multiple intensity images. Zuo *et al*. [44] introduced an electrically tunable lens and a 4*f* system into a conventional bright-field microscope and realized the non-mechanical focus control for TIE. In this technique, multiple captures of intensity images are still needed, but can be performed electrically, thus avoiding mechanical motion. Waller *et al*. [45] proposed the idea of combining volume holographic microscope with TIE to capture multiple intensity images in one frame. Though the single capture is achieved by this method, the image registration issue from the misalignment of several intensity images still exists.

1.3 Digital Holography

DH has been widely used in three-dimensional (3D) imaging and phase retrieval but needs a reference beam to first construct the hologram from interference of the object wave with the

reference [46-49]. Phase distributions of the object, or equivalently 3D profiles, can be reconstructed numerically by back propagation of the hologram function and subsequent phase unwrapping; however, issues such as the twin image problem during reconstruction need to be resolved. This is done either by off-axis holography, where a small angle is introduced between the object and reference beams [46-49], or through multiple registrations of the hologram in an in-line geometry where the phase of the reference is successively changed, as in phase-shifting DH (PSDH) [46-50]. Indeed, the principle of phase shifting to determine an unknown phase is general enough to be applied to other cases, viz., finding the phase, and thus, topography, of a 3D object using structured light [4].

1.4 Digital Holography with Transport of Intensity

DH and TIE can be combined to retrieve the phase from phase objects. For DH, phase unwrapping is needed to obtain the true phase values of the phase objects in the reconstruction process. Commonly used phase unwrapping algorithms like the phase unwrapping max-flow algorithm (PUMA) [51] suffer from some shortcomings. As shown in this thesis, phase unwrapping using PUMA can be time-consuming for large two-dimensional (2D) matrices, especially when phase excursions are large, typically in the range of hundreds of radians. Furthermore, we have observed that PUMA yields unsatisfactory results especially when the phase gradients are large, typically more than π radians per pixel. Now TIE, which yields the unwrapped phase, usually suffers from image registration issue for the moving components in the optical system. By combining DH with TIE, the phase unwrapping process in DH can be avoided, and image alignment can be easily achieved by numerically reconstructing the intensities at multiple planes around the image plane. Additionally, the unwrapped or true phase retrieval can be achieved with a single capture of the hologram.

Recently, Pandey *et al*. [52] proposed to unwrap a 2D wrapped phase map using TIE and obtaining intensities by numerical propagation. Martinez-Carranza *et al*. [53] developed a wave-

propagation free phase unwrapping algorithm based on TIE. Zhao *et al*. [54] proposed a robust phase unwrapping algorithm based on TIE using fast cosine transform. Single-shot phase imaging can also be achieved using randomized light [55], coded aperture [56], wavefront sensors [57], etc. Zuo *et al*. [58] proposed the idea of combining off-axis digital holography microscope (DHM) with TIE and showed experimental results on micro-bumps and human macrophage cells, with phase excursions in the order of tens of radians. Zhou *et al*. [59] combined in-line DH with TIE (DH+TIE) and performed experiments to find the phase across the cross-section of a tapered fiber. Phase excursions in this case was of the order of 3 radians. Both works demonstrate the advantages of combining DH with TIE, which are that there are no physical movements of sample or the detector, and no phase unwrapping is needed for determining the object phase. Zhou *et al*. [60] also performed simulations on off-axis DH+TIE, their conclusion states that the DH+TIE technique shows better reconstruction ability for selected complicated phase distributions with large phase excursions in theory, but no further experimental verification was performed.

1.5 Correlation of Holograms in 3D Recognition

Face, fingerprint, iris, voice are popular biometrics in human verification and identification. Verification is a 1-to-1 match that compares one person's feature with a template; while, identification usually refers to a 1-to-N problem by comparing one person's characteristics with other persons [61]. For both verification and identification, biometric techniques currently offer great accuracy and require cooperation from the users. For instance, 2D fingerprint acquisition requires physical contact with the sensor, and iris imaging requires users carefully matching their eye position with the sensor. The excessive reliance on the subject cooperation can sometimes exhibit difficulties in high-throughput applications [62]. In this situation, face recognition can be superior to the other biometric techniques. Two-dimensional face recognition by processing 2D intensity or color image has been reported a high recognition rate higher than 90% [61]. Even

though 2D face recognition has shown great potential, it still has limitations in illumination and drawbacks in security [63].

Three-dimensional human characteristics revealing total geometric shape offer much more details than 2D biometrics. For 3D fingerprint analysis, Abeywickrema *et al*. [6] proposed a topography approach for fingerprints by using multi-wavelength digital holography (MWDH). Based on the advancement of structured light scanning [64], the acquisition of 3D face data has been further realized. Several methods using 3D facial geometric information has been proposed and used in 3D face recognition. Blanz and Vetter [65] have proposed a 2D-based method for 3D face recognition. They estimate the shape and texture of a face by fitting a morphable model of 3D face to 2D images based on a vector representation. Then, comparison of two faces is achieved by Mahalanobis distances between the shape and texture parameters of the fitting model [66]. A typical 3D-based method is proposed by Ansari and Abdel-Mottaleb [67], which extracts feature points on eyes, noses, and mouths from frontal and profile images. Then the recognition is performed by calculating Euclidean distance between 29 features points. Multi-model approaches combining 2D face image and 3D face model have been investigated [68]. One method performs principle component analysis (PCA) on intensity and depth images and then synthesizes the results from both strategies to obtain a combined distance metric [69]. Recently, convolutional neural networks (CNNs) have become popular for 3D face recognition. The recognition can be achieved by designing layers, nodes, and parameters of the CNN architecture to evaluate normalized mean error (NME) of visible landmarks and mean average pixel error (MAPE) [70,71]. For all the methods described above, two pre-processed 3D faces are compared by distance or position metrics. All the 3D data are extracted or transformed into vectors or matrices; thus the information of the 3D data can be greatly reduced. Unfortunately, it is difficult to reconstruct the 3D data from the generated vector or matrix.

Optical parallel correlators have been used for 2D face recognition with high speed and accuracy [72]. Digital correlation of CGHs for identification has been first proposed by Abeywickrema *et al*. [73]. Zhou *et al*. [74,75] have extended correlation of CGHs to surface authentication. Their work reveals that CGHs are able to encode depth/height information of 3D objects into a 2D plane, and that 2D digital correlation of holograms can extract depth and intensity profile difference through evaluation of the correlation peak values. Matching of 3D objects by Wigner analysis of the correlation pattern between the phase-only holograms extracted using optical scanning holography as a form of electrical signal has been introduced by Kim and Poon [76].

1.6 Objectives and Significance of This Work

In the first part of this work, phase retrieval using TIE with TPE is proposed. A detailed description of the methodology is demonstrated. Simulations are performed to analyze the error, in order to justify the use of TPE with TIE. Experiments on onion epidermal cells and gratings are conducted to show the performance of TIE+TPE. TIE+TPE is a non-iterative phase retrieval technique that can yield fast performance using FFT and only three captures of the intensity images. This method can be developed to perform high-definition biomedical phase imaging.

The second part of this work is concerned with the retrieval of true unwrapped phase of phase objects using TIE. Both in-line and off-axis configurations of holographic recording are investigated. First, the twin image effect on phase retrieval using in-line DH+TIE is studied through simulations. The idea of combining PSDH with TIE (PSDH+TIE) is proposed through simulations. Thereafter, by defining suitable figures of merits, off-axis DH+TIE is compared with off-axis single-wavelength DH (SWDH) and dual-wavelength DH (DWDH) for different phase excursions, both in simulations and experiments. A computational performance analysis of DH+TIE has been performed and compared with standard DH methods involving retrieving unwrapped phase using PUMA, showing the superiority of DH+TIE in computational speed.

DH+TIE can achieve single-shot phase imaging without any phase unwrapping process, which can be potentially applied to real-time 3D imaging.

In the third part of this thesis, an idea of 3D face recognition by using 2D digital correlation of CGHs has been proposed. Identification and verification of 3D faces by using correlation of CGHs through three CGH algorithms for different propagation distances are demonstrated. The objective of this work is to achieve 3D face recognition in a 2D correlation scenario. By adjusting the size and pixel number of the CGH, the computational time can be reduced. A 2D hologram used as a face ID can be secure in privacy protection through encoding depth into 2D plane. Real 3D face features can be reconstructed and displayed by backward propagation of holograms with potential applications in optical pattern recognition [77] and 3D visualization [78]. The technique can also be applied to digitally recorded holograms. This is beneficial in identifying 3D surface features, e.g. the surfaces of additively manufactured objects.

1.7 Organization of the Thesis

The organization of the thesis is as follows. Chapter II introduces TIE phase imaging with TPE error correction method, along with simulations and experiments. Chapter III shows the performance analysis using DH+TIE with different geometries. Chapter IV gives the implementation of correlation of holograms in face recognition. In the last Chapter, concluding remarks are presented and future work directions are pointed out.

1.8 Publications

A number of refereed journal articles, along with conference proceedings and conference presentations have been published while working at the Department of Electro-Optics and Photonics, University of Dayton:

[Journal papers]

1. H. Zhou, E. Stoykova, M. Hussain, and P. P. Banerjee, "Performance analysis of phase retrieval using transport of intensity with digital holography," Appl. Opt. **60**, A73-A83 (2020).
2. H. Guo, H. Zhou, and P. P. Banerjee, "Single-shot digital phase-shifting Moiré patterns for 3D topography," Appl. Opt. **60**, A84-A92 (2020).
3. H. Zhou, X. Sui, L. Cao, and P. P. Banerjee, "Digital correlation of computer-generated holograms for 3D face recognition," Appl. Opt. **58**, G177-G186 (2019).
4. B. Bordbar, H. Zhou, P. P. Banerjee, "3D object recognition through processing of 2D holograms," Appl. Opt. **58**, G197-G203 (2019).

[Conference presentations / proceedings]

1. H. Zhou and P. P. Banerjee, "Transport of intensity phase imaging with error correction using transport of phase equation," Proc. SPIE **11709**, 117090D (2021).
2. H. Zhou, E. Stoykova, and P.P. Banerjee, "Phase retrieval using transport of intensity with off-axis digital holography for objects with large phase excursions", HF2D.5, Digital Holography and 3D Imaging, OSA (2020).
3. E. Stoykova, H. Zhou, and P.P. Banerjee, "Phase retrieval by transport of intensity in inline digital holography", HF2D.3, Digital Holography and 3D Imaging, OSA (2020).
4. H. Guo, H. Zhou, and P. P. Banerjee, "Single-shot Digital Phase-shifting Moiré Pattern for 3D Metallic Surface Imaging," HF3G.3, Digital Holography and 3D Imaging, OSA (2020).
5. H. Gao, H. Fang, J. Liu, H. Zhou, X. Cheng, S. Ding, J. Luo, S. Li, Z. Dai, and P.P. Banerjee, "A scanning method based on parabolic mirror and galvanometer for holographic contact copying," HTh4H.1, Digital Holography and 3D Imaging, OSA (2020).

6. H. Zhou, R. Hou, B. Bordbar, and P. P. Banerjee, “Effect of hologram windowing on correlation of 3D objects,” Th2B.8, Digital Holography and 3D Imaging, OSA (2019).

7. H. Zhou, R. Hou, B. Bordbar, and P. P. Banerjee, “Effect of hologram size on 3D reconstruction using multi-wavelength digital holography,” W4B.2, Digital Holography and 3D Imaging, OSA (2019).

8. P. P. Banerjee, U. Abeywickrema, H. Zhou, M. S. Alam, G. Nehmetallah, J. Khoury, L. Cao, “Taking correlation from 2D to 3D: optical methods and performance evaluation,” Proc. SPIE **10995**, 10995-10 (2019).

9. H. Zhou, U. Abeywickrema, B. Bordbar, L. Cao, P. P. Banerjee, “Correlation of holograms for surface characterization for diffuse objects,” Proc. SPIE **10943**, 10943-3 (2019).

CHAPTER II

PHASE RETRIEVAL USING TRANSPORT OF INTENSITY WITH TRANSPORT OF PHASE

As discussed in Chapter I, TIE can be solved to directly retrieve phase. In this Chapter, a non-iterative way of solving TIE with the help of TPE is proposed. The effect of diffraction, which is embedded in TPE, will be included to find a more accurate phase solution compared to the FFT-based TIE method. Simulations of Gaussian phase are performed to show the feasibility of the method. Experiments on phase objects such as onion cells and a grating are presented to analyze the performance of the proposed method.

2.1 Methodology

The TIE+TPE method can be considered as a two-step phase retrieval process. First, the FFT method is used to retrieve the phase under the premise that intensity at focused or image plane is a constant. In the second step, a correction factor is introduced to compensate the errors from finite difference of longitudinal derivatives of intensity and the nonuniformity of the intensity at focused plane. TPE is used to calculate the longitudinal phase derivative in the correction factor. Then, TIE is solved again with the correction factor using, once again, an FFT-based method.

The starting point is the paraxial form of TIE (detailed derivation in Appendix A):

$$\boldsymbol{\nabla}_{\perp} \cdot (I \boldsymbol{\nabla}_{\perp} \phi) = -k_0 \frac{\partial I}{\partial z}, \tag{2-1}$$

where I is the intensity of the optical field, ϕ is the phase of the envelope of the optical field, $\boldsymbol{\nabla}_{\perp}$ is the transverse gradient operator, $k_0 = \frac{2\pi}{\lambda}$ is the wave number in free space, λ is the wavelength. Under the assumption that the intensity at the focal plane is approximately a uniform distribution,

the intensity term in the left-hand side (LHS) of Eq. (2-1) can be moved out of the $\boldsymbol{\nabla}_\perp$ operator. Then, the TIE is reduced to a Poisson equation

$$\nabla_\perp{}^2\phi + k_0 \frac{1}{I}\frac{\partial I}{\partial z} = 0\,, \tag{2-2}$$

which can be solved by an FFT method. The solution of the phase for Eq. (2-2) is:

$$\phi = \Im^{-1}\left\{\frac{1}{k_x^2+k_y^2}\Im\left\{\frac{k_0}{\bar{I}}\frac{\partial I}{\partial z}\right\}\right\}. \tag{2-3}$$

where k_x and k_y are the transverse spatial frequencies, $\bar{I}$ is the intensity at the image (also called the focused) plane and can be taken as the average intensities at two symmetrically defocused planes [79]. To calculate the longitudinal derivative of the intensity, a finite difference method can be used, assuming the high-order derivative terms can be negligible. Then the derivative in Eq. (2-3) can be re-expressed as

$$\frac{\partial I}{\partial z} \approx \frac{I(+\Delta z)-I(-\Delta z)}{2\Delta z}. \tag{2-4}$$

The assumptions made in solving the TIE with the FFT method stated above limit the accuracy of the TIE. Although the TIE is mostly used for phase objects to determine the focused phase, errors can be introduced in the phase retrieval result if the intensity is not ideally uniform. Higher-order errors in the finite differencing of the derivative $\frac{\partial I}{\partial z}$ in Eq. (2-4) can be compensated by using more captures in TIE imaging [80]; however, this may require a large stack of intensity images along the z-axis.

To maintain the exactness of Eq. (2-1) but express it as a Poisson-type equation as in (2-2), one can introduce a function $n(x, y)$ such that

$$I(\nabla_\perp^2\phi) + k_0 n\frac{\partial I}{\partial z} \equiv \boldsymbol{\nabla}_\perp \cdot (I\boldsymbol{\nabla}_\perp\phi) + k_0\frac{\partial I}{\partial z} = 0. \tag{2-5}$$

Then the correction factor can be expressed as

$$n = 1 + \frac{1}{k_0 \frac{\partial I}{\partial z}} \boldsymbol{\nabla}_\perp I \cdot \boldsymbol{\nabla}_\perp \phi + \frac{1}{k_0} \frac{\partial \phi}{\partial z}, \tag{2-6}$$

which can be thought of as an effective inhomogeneous medium with a refractive index $n(x, y)$. To calculate $\frac{\partial \phi}{\partial z}$, the TPE is employed. A general non-paraxial expression of the TPE is

$$\boldsymbol{\nabla}\tilde{\phi} \cdot \boldsymbol{\nabla}\tilde{\phi} = k_0^2 + \frac{1}{\sqrt{I}} \nabla^2 \sqrt{I}, \tag{2-7}$$

where $\tilde{\phi} = \phi + k_0 z$ denotes the phase of the phasor component of the optical field and $\boldsymbol{\nabla}$ is the three-dimensional gradient operator. Since both the TIE and TPE are supposed to be satisfied simultaneously, the captured intensity and computed phase from Eq. (2-3) should satisfy the TPE at the focused plane. Note that if the intensity is ideally uniform, or equivalently in the absence of diffraction [27], Eq. (2-5) is simplified to

$$\boldsymbol{\nabla}\tilde{\phi} \cdot \boldsymbol{\nabla}\tilde{\phi} = k_0^2. \tag{2-8}$$

However, this condition is never true in any physical situation. From Eq. (2-7), the derivative of the phase ϕ *w.r.t.* the longitudinal direction z can be written as:

$$\frac{\partial \phi}{\partial z} = \sqrt{k_0^2 + \frac{1}{\sqrt{I}} {\nabla_\perp}^2 \sqrt{I} + \frac{1}{\sqrt{I}} \frac{\partial^2 \sqrt{I}}{\partial z^2} - \boldsymbol{\nabla}_\perp \phi \cdot \boldsymbol{\nabla}_\perp \phi} - k_0. \tag{2-9}$$

With Eqs. (2-6) and (2-9), Eq. (2-5) can be solved again to obtain the phase with correction. The result, similar to Eq. (2-3), can be expressed as:

$$\phi = \Im^{-1} \left\{ \frac{1}{k_x^2 + k_y^2} \Im \left\{ \frac{k_0 n(x,y)}{\bar{I}} \frac{\partial I}{\partial z} \right\} \right\}. \tag{2-10}$$

As seen from Eq. (2-10), the "virtual" inhomogeneous medium now has a *local* propagation constant $k(x, y) = k_0 n(x, y)$ and is still a Poisson-type equation. The solution method is still based on FFT; thus, it is easy to implement and efficient in computation. The defocused intensity images in the simulations are generated by using the non-paraxial diffraction formula. Since the optical fields need to be propagated in a very small distance (micro-scale) [81], the distances from

the defocused intensity image planes to the focused (image) plane are set as $\pm\Delta z$. The (nonparaxial) impulse response of the propagation of light can be expressed as [59]:

$$h_p(x, y; \Delta z) = \frac{jk_0}{2\pi} \frac{exp\left[-jk_0\sqrt{x^2+y^2+(\Delta z)^2}\right]}{\sqrt{x^2+y^2+(\Delta z)^2}}. \tag{2-11a}$$

The corresponding nonparaxial transfer function is:

$$H_p(k_x, k_y; \Delta z) = \exp\left[-j\left(\sqrt{{k_0}^2 - ({k_x}^2 + {k_y}^2)}\Delta z\right)\right], \tag{2-11b}$$

where k_x and k_y are the lateral spatial frequencies. The paraxial versions of the impulse response and transfer function can be readily found by approximating the square roots in the above expressions.

2.2 Error Source Analysis and Simulations

In the FFT-based TIE method, the two major sources of errors considered are from

(a) the finite difference estimation of the intensity at the focused or image plane;

(b) the nonuniformity of the intensities and exclusion of the diffraction effect.

To analyze TIE+TPE method's correction for these errors, simulations for Gaussian phase distributions with uniform and non-uniform intensity profiles are conducted and results are shown below.

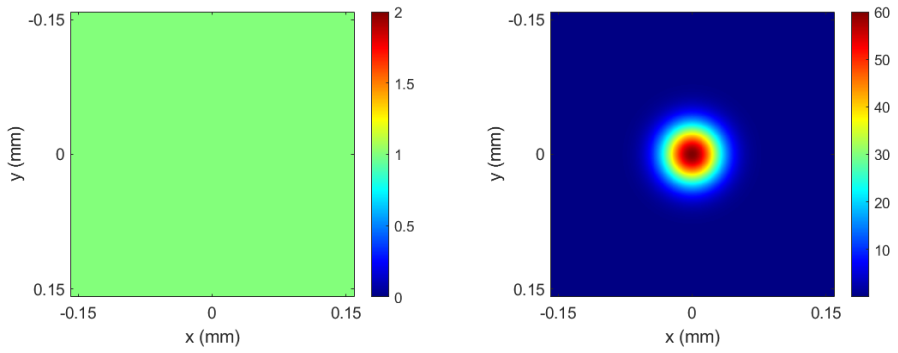
2.2.1 Object with Uniform Intensity

Starting from an ideal phase object, the intensity at the focused plane is set to be a uniform distribution (Figure 2-1(a)). The FFT-based TIE method should give an almost exact solution of the phase profile, with possible minor errors coming from the finite difference approximation to the first derivative in longitudinal direction stemming from high-order derivative terms. Under the

condition of constancy of the intensity in transverse plane, $\boldsymbol{\nabla}_\perp I = 0$ and $\frac{1}{\sqrt{I}}\nabla_\perp{}^2\sqrt{I} = 0$. Then, the correction factor in Eq. (2-6) with Eq. (2-9) can be simplified as

$$n(x,y) = \sqrt{1 + \frac{1}{\sqrt{I}}\frac{\partial^2\sqrt{I}}{\partial z^2} - \frac{1}{k_0}\boldsymbol{\nabla}_\perp\phi \cdot \boldsymbol{\nabla}_\perp\phi}. \tag{2-12}$$

In what follows, we first test the effectiveness of using TIE and TPE through the usage of Eq. (2-12) with Eq. (2-10) in reducing the errors in the finite difference estimation. To this end, a numerical verification is performed using a phase object with Gaussian phase distribution and a uniform intensity at focused plane $z = 0$. The peak of the Gaussian phase is 60 rad with a width of 50 µm. This is shown in Figure 2-1(b). The object is designed with a size of 1024×1024 pixels with a pixel pitch of 0.31 µm. For the simulation, a coherent illumination at a wavelength of 660 nm is used. All the parameters are selected based on experimental parameters. The TIE+TPE method is used to retrieve the unwrapped phase at $z = 0$.



(a) (b)

Figure 2-1: The (a) intensity and (b) phase of the designed Gaussian phase object for simulation. The intensity is designed as a uniform distribution, which indicates the object is an ideal phase object. The maximum phase value of the Gaussian phase is 60 rad with a width of 50 µm.

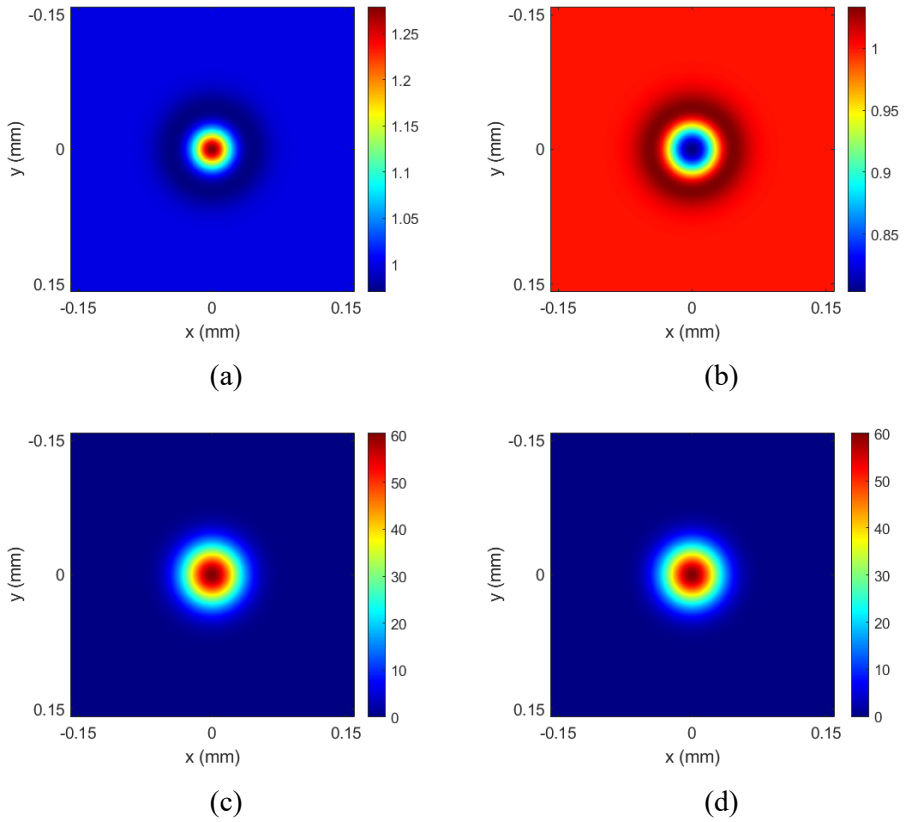


Figure 2-2: The numerical generated intensity images at the planes $z =$ (a) $-\Delta z$; (b) Δz. Phase retrieval results using (c) FFT-based TIE; (d) TIE+TPE for the Gaussian phase object designed in Figure 2-1.

Assuming the object is located at $z = 0$, the object field is propagated to $z = \pm\Delta z$, where Δz is set to 10 μm. By computing the convolution between the object field and the impulse response in Eq. (2-11a), the optical field $\Gamma(x, y; z)$ at defocused planes can be obtained. The convolution can be achieved by taking the Fourier transform of the object field and the impulse response, multiplying them in the spatial frequency domain, and performing the inverse Fourier Transform for the multiplied field. Then, the intensity can be calculated by $I(x, y; z) = |\Gamma(x, y; z)|^2$. The intensity images at the defocused plane are shown in Figures 2-2(a,b). For simplicity, we define $I_a = I(x, y; a\Delta z)$, $a = -1,0,1$. From the intensity images, the phase distribution at the focused plane can be computed using the FFT-based TIE (see Eq. (2-3)), as shown in Figure 2-2(c). Then, the intensity images and calculated phase distribution are used to calculate the correction factor in

Eq. (2-10), using Eq. (2-12) in this case. The first-order derivative of phase at the focal plane is computed in transverse dimensions, respectively. Once the correction factor $n(x, y)$ is obtained, the TIE is solved by using the updated TIE+TPE equation (Eq. (2-12)). The solution of the original TIE is shown in Figure 2-2(c); the TIE+TPE result is presented in Figure 2-2(d). The small DC components (~0.1 rad) in the results have been removed for better comparison. Though the difference in both results may not be readily obvious, the subtle difference can be detected by using a performance metric such as root mean squared error (RMSE), defined as

$$RMSE = \sqrt{\frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} \left[\phi_{retrieved}(i,j) - \phi_{ground\ truth}(i,j)\right]^2}\ . \tag{2-13}$$

A small value of RMSE indicates that the retrieved phase is close to the ground truth. The RMSEs for the Gaussian phase with a maximum value of 60 rad is calculated for FFT-based TIE and TIE+TPE by comparing with the ground truth of the Gaussian phase profile. The result of RMSE for the FFT-based TIE is 5.49×10^{-2} rad; while for the TIE+TPE is 4.25×10^{-2} rad. Both are very small numbers and close to each other; however, the TIE+TPE shows a 22.6% reduction in error over the conventional TIE. Since the only error here should be from finite differencing, it is evident that this has indeed been reduced using the TIE+TPE, thereby establishing the justification of the TIE+TPE method.

2.2.2 Object with Non-uniform Intensity

To check the performance of the TIE+TPE in the presence of non-uniform intensity, simulations are performed as follows. The same Gaussian phase distribution in Figure 2-1(b) is used. As stated earlier, in practice, it is hard to achieve a constant intensity. Furthermore, even in the Gaussian phase case in Section 2.2.1, a slight displacement from the focused plane introduces a Gaussian-like intensity profile. To examine the efficacy of the TIE+TPE technique, we have assumed a Gaussian profile with a maximum amplitude value of 0.01 and a width of 50 μm (same as the width of the phase profile) added to the constant intensity at the focused plane, as shown in

Figure 2-3(a). A similar correction process using the TIE+TPE as described above is also performed with this complex object having the phase from Figure 2-1(b) and intensity from Figure 2-3(a). The phase retrieval results by using FFT-based TIE and TIE+TPE are shown in Figure 2-3(b,c) respectively. To show the difference in detail, a cut along the *x*-axis across the center of the phase profile is extracted and plotted in Figure 2-3(d). The black line is the ground truth; the blue line is the result from the TIE+TPE method; the red line is obtained from FFT-based TIE result. The insets in Figure 2-3(d) clearly show the improvement using TIE+TPE.

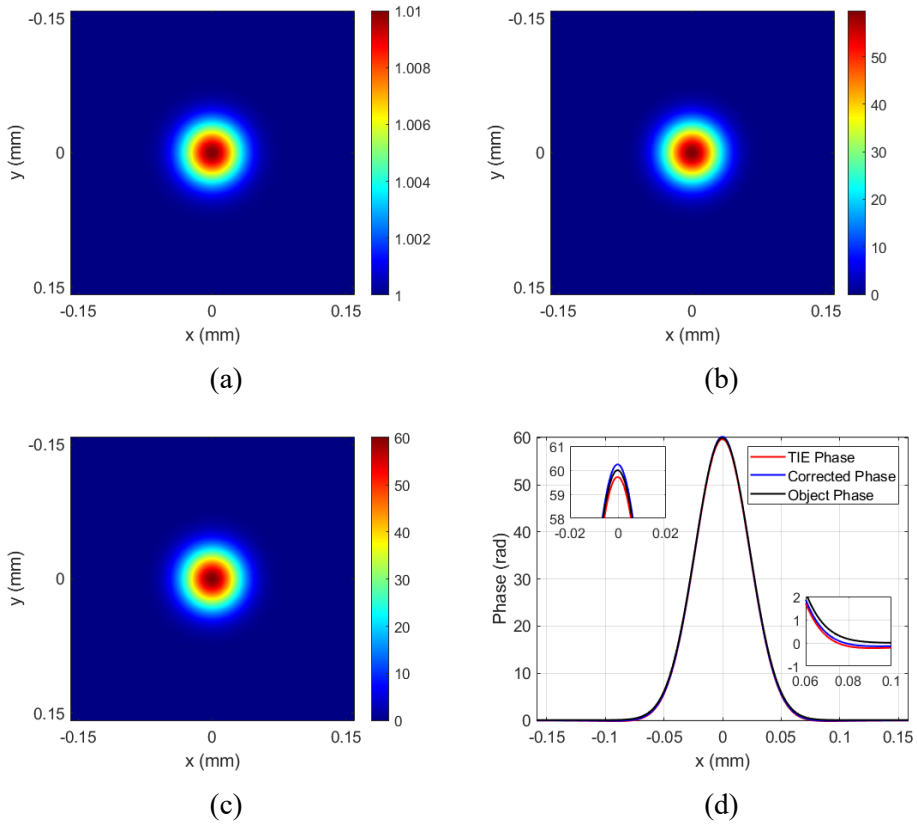


Figure 2-3: (a) Non-uniform intensity profile of the designed object in evaluating the performance of TIE+TPE. Phase retrieval results using (b) FFT-based TIE; (c) FFT-based TIE+TPE for object with phase in Figure 2-1(b) and intensity in (a). (d) Extracted profiles across the centers of (b,c) along *x*-axis. The insets are the zoomed-in plots of the Gaussian phase peak and the right edge of the Gaussian phase peak.

Once again, the RMSEs are calculated with the ground truth of the phase profile: for the FFT-based TIE, it is 0.249 rad; while for the TIE+TPE method, it is 0.156 rad. Though the performance

for both methods degraded as compared to the phase object case, the TIE+TPE gives a more accurate result than traditional TIE method.

To further investigate TIE+TPE method performance with different perturbations in the intensity, the Gaussian intensity perturbation with amplitude of 0.01 to 0.07 in steps of 0.02 are applied to the constant intensity. A similar process is done for all the cases. The RMSEs are calculated and presented in Table 2-1. From Table 2-1, it is clear that as the amplitude of the intensity perturbation increases, the RMSEs of the phase retrieval results for both methods will increase; however, the TIE+TPE method still outperforms the FFT-based TIE method.

Table 2-1: RMSEs for TIE and TIE+TPE phase retrieval with different intensity perturbation

RMSE (rad)	Intensity perturbation amplitude			
	0.01	0.03	0.05	0.07
FFT-based TIE	0.249	0.794	1.33	1.86
TIE+TPE	0.156	0.403	0.723	0.945

2.3 Experimental Results

2.3.1 Experimental Setup

Experiments are conducted with a commercial Nikon T*i*-E series inverted microscope, as shown in Figure 2-4. A halogen lamp (12V 100 W) is used as the light source in the experiment, which provides a broad-band incoherent light. The Kohler illumination setup is used in the microscope. Light goes through filters and infrared region is blocked (> 800 nm). The two diaphragms are used to control how much light will illuminate on the sample. After light focuses on the sample, located on a 3D translational stage, it is collected by the microscope objective 10×/0.30 NA. An intermediate image will form at the rear focal plane of the tube lens system. The intermediate image plane is indicated in Figure 2-4 with green dashed lines. Finally, the intensity

image of the sample is formed at the CCD camera plane. Note that the tube lens system consists of a Galilean telescope (a positive lens followed by a negative lens) with a magnification of 1.5 × and a single tube lens.

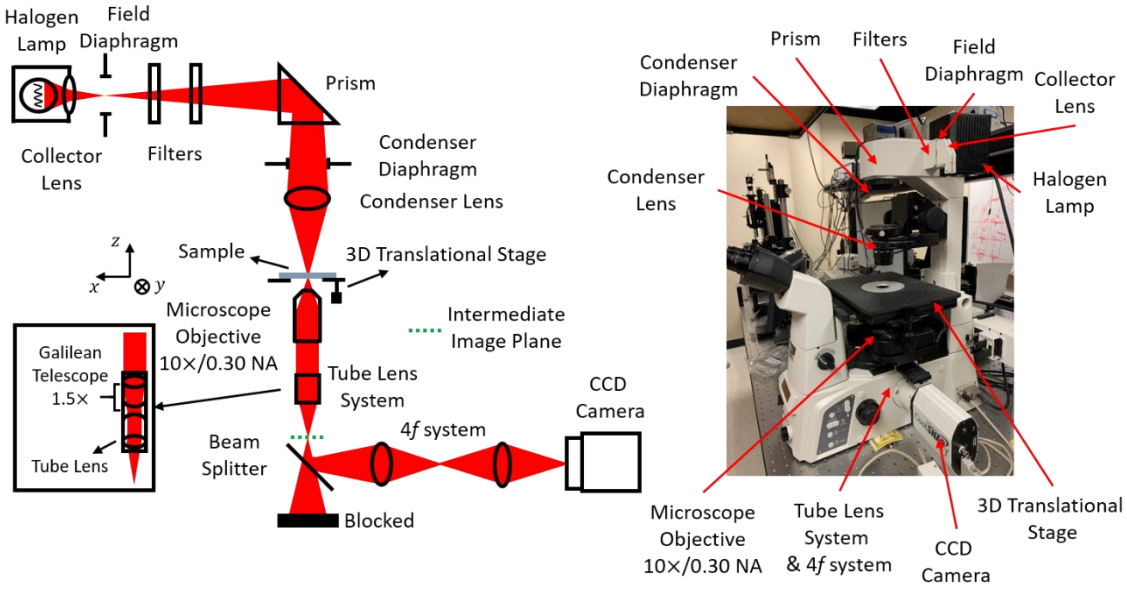


Figure 2-4: Schematic (left) and photograph (right) of the experimental setup for TIE and TIE+TPE phase retrieval experiments. The halogen lamp (12V 100W) is the incoherent source. The Kohler illumination is used to avoid filament of the lamp from forming an image at the CCD plane. The filters are used to filter out infrared light with a wavelength longer than 800 nm. The 3D translational stage can help select the region of interest and achieve 50 nm precision movement along light propagation direction. The CCD camera has 1024×1024 pixels. The pixel pitch after calibration is 0.31 μm/pix. The green dashed line is the intermediate image plane.

The movements of the 3D translational stage along the transverse x- and y-directions are mechanically driven; while along the longitudinal or z-axis is piezo-electric driven, which can achieve resolutions of 50 nm per step. By moving the sample incrementally along the z-axis, defocused intensity images can be captured and used for phase retrieval using TIE and TIE+TPE methods. The microscope system has a magnification of 15×, with a magnification of 10× from microscope objective and a magnification of 1.5× from an intermediate magnification changer (also known as optovar). The CCD camera has a pixel pitch of 4.65 μm/pix. After system calibration, the effective pixel pitch for the original intensity images is 0.31 μm/pix. Each

recorded intensity image is 1024×1024 pixels. The (free space) wavelength of the incoherent source, averaged over the power of the spectrum of the light source (Figure 2-5), is $\lambda = 659.5$ nm. This wavelength is used to convert phase radians into height/depth values using the relation $\phi = \frac{2\pi n}{\lambda} z$, where n is the refractive index of the transmissive sample.

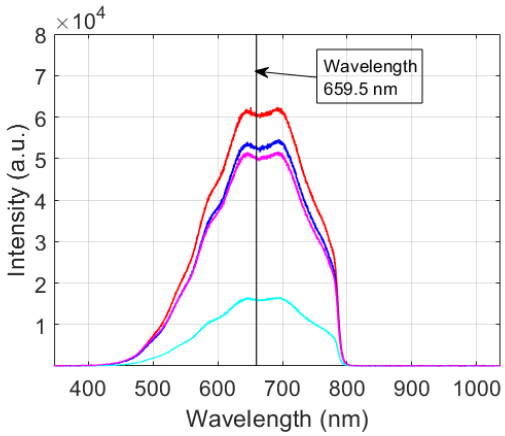
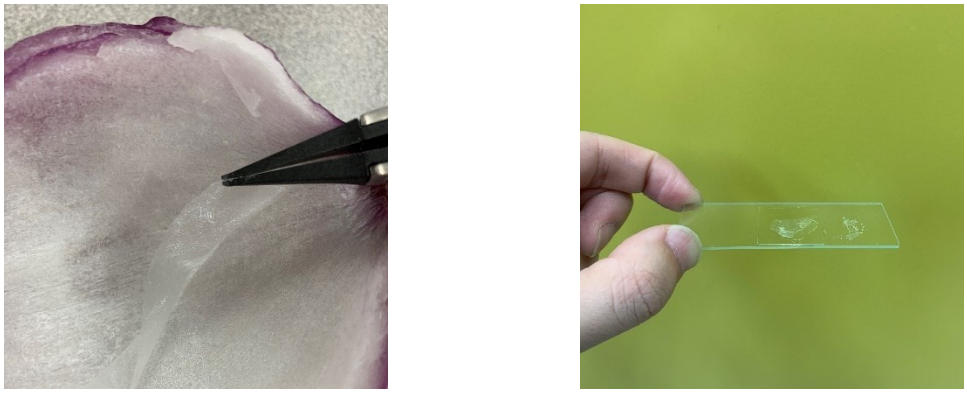


Figure 2-5: Spectrum of the halogen light source after filtering, measured at the sample plane. Different colors denote spectra with different sizes of the diaphragm. The black line in the inset is the averaged wavelength weighted by the power of the spectrum for all measurements, which is 659.5 nm.

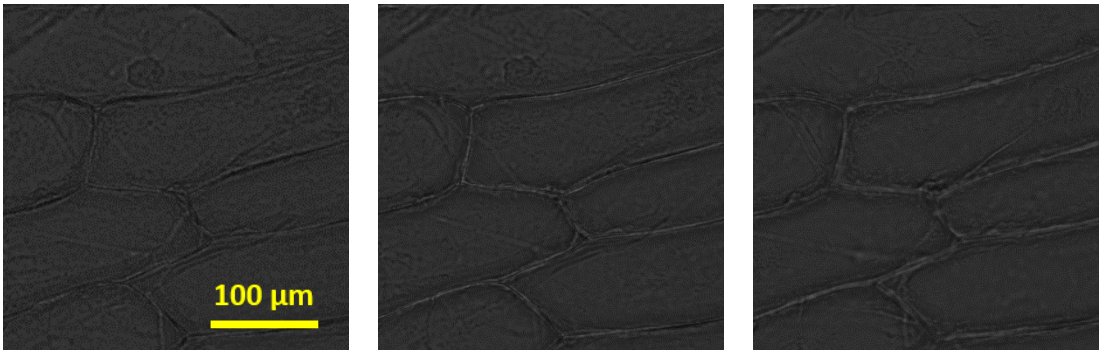
2.3.2 Phase Retrieval of Onion Epidermal Cells



(a) (b)

Figure 2-6: (a) Onion epidermal cells preparation. A tweezer is used to tear the epidermal layer off the onion. (b) Onion epidermal cells in saturated salt water on a glass slide.

Cells are common phase objects in our daily life. Here, onion epidermal cells are used as a transmissive phase object for the verification of TIE+TPE method. The epidermal cells of onions provide plant tissues protection from viruses. A single layer of onion epidermal cells can be obtained by using a tweezer to tear it off the onion (Figure 2-6(a)). Then, the layer of onion epidermal cells is cut into two small pieces and put on a glass slide (Figure 2-6(b)). One piece of the cell sample is placed in purified water for one minute; the other one is placed in a saturated salt solution (at room temperature) for one minute. Then, the samples are illuminated by a halogen lamp and imaged by a microscope, as described above (see Figure 2-4 for details). The intensity images at focused and defocused planes are captured (Figure 2-7).



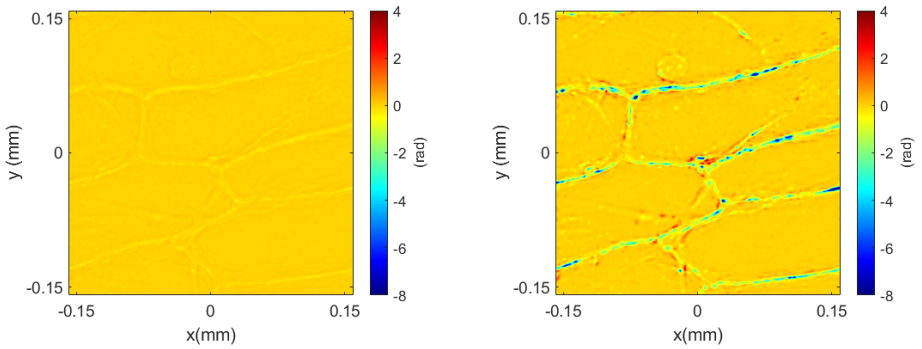
(a) (b) (c)

Figure 2-7: Intensity images of onion epidermal cells after immersion in water for one minute at (a) $z = -\Delta z$; (b) $z = 0$; (c) $z = \Delta z$; $\Delta z = 10$ µm.

From the three intensity images located at different planes along z-axis, the phase profiles can be retrieved by FFT-based TIE using Eq. (2-3) (Figure 2-8(a)) and TIE+TPE using Eq. (2-10) (Figure 2-8(b)). The colors in Figure 2-8 indicate the scale of possible variations of the phase profile ranging from -8 rad to 4 rad. According to Agarwal *et al*. [82], the phase values of cell walls are smaller than the phase values of the inner cell, and the phase variations range from -8 rad to 4 rad at a wavelength of 632.8 nm using structured light illumination exploiting the Talbot effect. However, from Figure 2-8(a), the phase variations retrieved by the FFT-based TIE are much less than the values reported in Agarwal *et al*. [82]. In the TIE+TPE result, a clear boundary of the cell

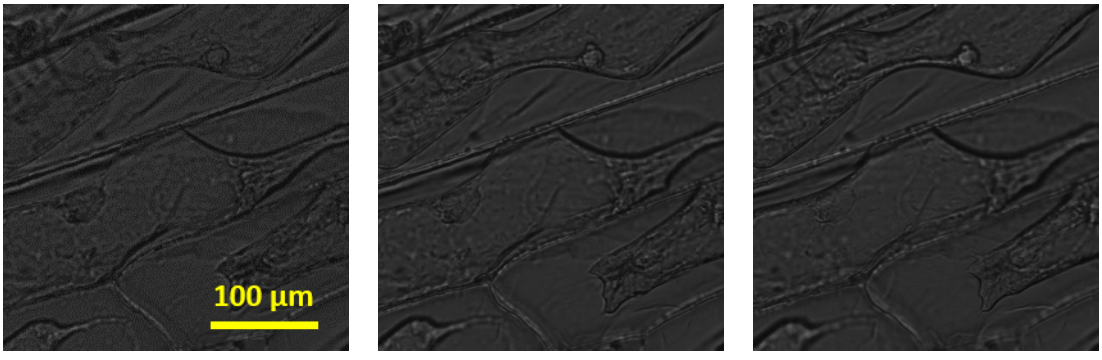
walls can be observed (green and blue sections in Figure 2-8(b)) along with a more marked phase variation, which also matches the result in Agarwal *et al*. [82].

The other onion epidermal sample is placed in saturated salt water. The three captures of intensity images are shown in Figure 2-9 and the phase retrieval results using FFT-based TIE method and TIE+TPE method are presented in Figure 2-10(a,b), respectively. Compared to the result in Figure 2-10(a), improvements in correcting the phase values from TIE+TPE in Figure 2-10(b) can be observed. Figure 2-10(b), shows the plasmolysis effect where plant cells lose water in a hypertonic solution and the cell wall remains its original structure. The green and blue sections indicate the cell walls. The cell plasma membranes shrink (reddish edges in Figure 2-10(b)) and are detached from the cell walls (greenish and blueish edges in Figure 2-10(b)).



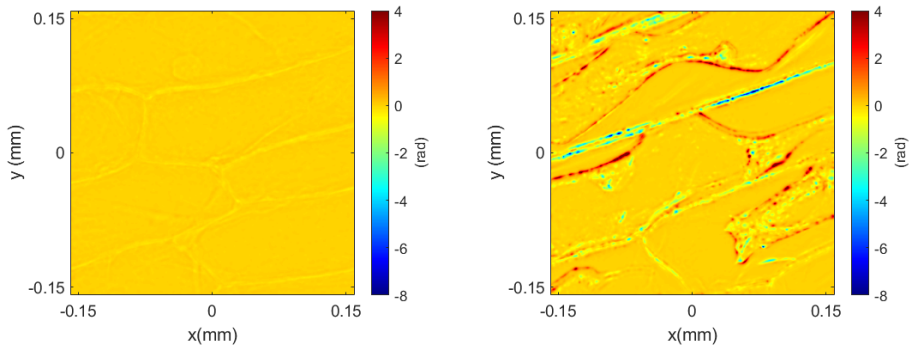
(a) (b)

Figure 2-8: Phase retrieval profiles for onion epidermal cells in purified water using (a) FFT-based TIE; (b) TIE+TPE. Colormap indicates the phase variations of the sample using $\lambda = 659.5$ nm.



(a) (b) (c)

Figure 2-9: Intensity images of onion epidermal cells in saturated salt water solution for one minute at (a) $z = -\Delta z$; (b) $z = 0$; (c) $z = \Delta z$; $\Delta z = 10$ µm.

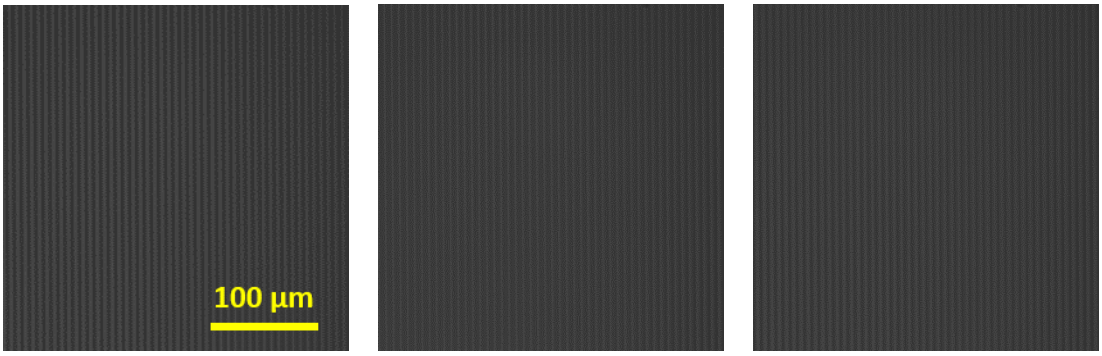


(a) (b)

Figure 2-10: Phase retrieval profiles for onion epidermal cells in a saturated salt water solution using (a) FFT-based TIE; (b) TIE+TPE. Colormap indicates the phase variations of the sample using $\lambda = 659.5$ nm.

2.3.3 Topography Retrieval of Gratings

In this example, imaging of a grating using the setup in Figure 2-4 is performed by using FFT-based TIE and TIE+TPE. The grating is made of glass with a period of 6 μm, 50% duty cycle, and with etching depth of 100 nm. The three intensity images are shown in Figure 2-11. Phase retrieval by FFT-based TIE and TIE+TPE are presented in Figure 2-12(a,b), respectively, with phase values converted into depth information (in nm). The grating depth profile is designed to be rectangular: a line scan by a profilometer yields the profile as shown in Figure 2-12(c). Though the scanning result from profilometer may not be very accurate, it still gives an approximately value of the depth. Comparison of Figures 2-12(a,b) with Figure 2-12(c) shows that our TIE+TPE approach shows a closer agreement to the profilometer scan compared to the FFT-based TIE result. It is encouraging to note that 3D surface mapping of 100 nm scale depths is resolvable using incoherent illumination and TIE+TPE. This may offer obvious advantages over white light interferometry; and will be pursued in the future.



(a) (b) (c)

Figure 2-11: Intensity images of one dimensional (1D) grating with a period of 6 µm, a depth of 100 nm, and a duty cycle of 50% at (a) $z = -\Delta z$; (b) $z = 0$; (c) $z = \Delta z$; $\Delta z = 10$ µm.

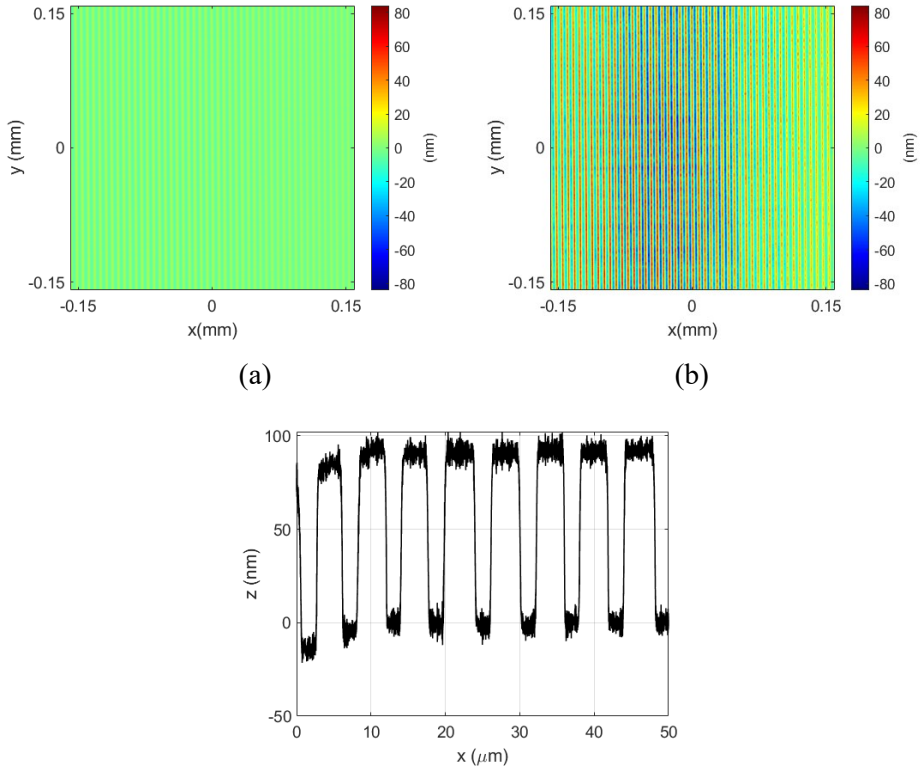


Figure 2-12: Phase retrieval profiles for 1D grating with a period of 6 µm, a depth of 100 nm, and a duty cycle of 50% using (a) FFT-based TIE; (b) TIE+TPE. Color map indicates the depth information (unit in nm) of the sample using mean wavelength $\lambda = 659.5$ nm. (c) A 1D scan of the grating using profilometer.

2.4 Conclusion

In this Chapter, a non-iterative FFT-based TIE solution with error correction using the TPE is proposed. A detailed description of the method is described. Based on the commonly used FFT-based TIE phase retrieval, a correction factor is introduced to compensate for the errors from the high-order derivatives from the finite difference in the estimation of longitudinal intensity derivation and the non-uniformity of the intensity. The correction factor can be perceived as a local change of the refractive index: the physical concept is to compensate errors by assuming that light is traveling through a virtual inhomogeneous medium. The computation of the correction factor is achieved by incorporating the TPE. Furthermore, by using the TPE, the diffraction effect is automatically taken into account.

Simulations are conducted to evaluate the performance of TIE+TPE in correcting expected errors. From monitoring RMSEs, the improvements can be observed by comparing FFT-based TIE result for both ideal phase object (constant intensity) case and phase objects with intensity perturbations. In experiments, onion epidermal cells in purified water and saturated salt water solution are studied. TIE+TPE phase retrieve results show a better accuracy than FFT-based TIE method. The same conclusion can also be reached in the 3D profilometry for gratings. Experiments are done using incoherent light, and attests to the versatility of the TIE and TIE+TPE technique.

In the next Chapter, the TIE will be combined with digital holography to retrieve phases from lenses.

CHAPTER III

PERFORMANCE ANALYSIS OF PHASE RETRIEVAL USING DH+TIE

The reconstruction of the object, particularly phase objects, using DH requires back/forward propagation of the hologram function using the Fresnel diffraction formula or the transfer function for propagation. However, the reconstructed phase is wrapped and, particularly for large phase excursions, can cause errors during unwrapping using PUMA. In this Chapter, TIE is combined with DH to directly retrieve the unwrapped object phase from a single capture of the intensity, in this case, the hologram. Multiple registrations of the intensities around the focus or image plane required for TIE are achieved numerically, thereby eliminating alignment issues encountered in an actual measurement. Performance of the DH+TIE in different geometries are analyzed and compared. Both simulations and experiments are conducted and show consistent results.

3.1 Theory and Method of DH+TIE

The general idea of DH+TIE consists of three steps:

(a) recording the digital hologram from the object,

(b) reconstructing intensity images at the image plane and two slightly defocused planes,

(c) employing TIE using these reconstructed intensity images to retrieve the unwrapped object phase.

3.1.1 Recording of Holograms

The hologram recording process involves interfering the diffracted object field $E_{od}(x,y) = E_o(x,y;z=d)$ from the initial (phase) object $E_{o0}(x,y) = E_o(x,y;z=0) = A_o e^{-j\phi_o(x,y)}$ at (see Figure 3-1) with a (plane wave) reference beam $E_r(x,y;z=d) = A_r e^{-j\phi_r}\, e^{-j(k_{0x}x+k_{0y}y)}$ at the hologram (or detector) $z=d$. It is assumed that A_r and ϕ_r are constants, as is A_o (for an ideal

phase object) and $\phi_o(x,y)$ denotes the object phase. Also $k_{0x} = k_0 \sin(\theta_x)$, $k_{0y} = k_0 \sin(\theta_y)$, where θ_x and θ_y are the tilt angles along x- and y-axes, respectively, and $k_0 = \frac{2\pi}{\lambda}$ is the wave number, where λ is the wavelength of the light. For an in-line hologram, $\theta_x = \theta_y = 0°$. The diffracted object field is related to the initial object field through the diffraction formula

$$E_{od}(x,y) = E_o(x,y;z=d) = E_o(x,y;0) * h_p(x,y;d) = E_{o0}(x,y) * h_p(x,y;d), \quad (3\text{-}1)$$

where $h_p(x,y;z)$ is the impulse response of propagation [59]:

$$h_p(x,y;z) \propto \frac{e^{-jk_0\sqrt{x^2+y^2+z^2}}}{\sqrt{x^2+y^2+z^2}}. \quad (3\text{-}2)$$

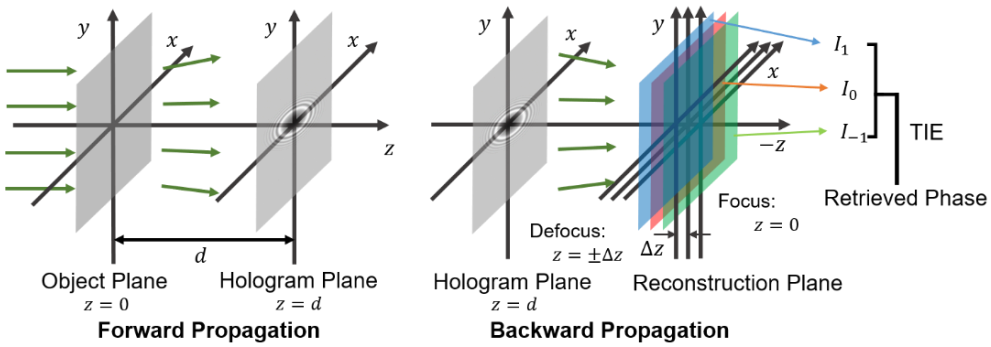


Figure 3-1: A schematic diagram of the DH+TIE method. The green arrows show the propagation direction of laser beams. The distance from the object plane to the hologram plane is d. Backward propagation process is from hologram plane to the reconstruction plane (the red plane) which is also located at $z = 0$. The blue and green planes are slight defocused from the red plane with a separation of Δz. From each of the planes, the intensity images of the object field can be acquired, and these intensity images (viz. I_1, I_0, I_{-1}) can be used in TIE to retrieve the phase of the object directly. [79]

The hologram function, which is proportional to the intensity pattern resulting from the interference of the object and reference fields, is given by

$$h_{\phi_r}(x,y;d) = \left|E_{od}(x,y) + A_r e^{-j\phi_r} e^{-j(k_{0x}x+k_{0y}y)}\right|^2$$

$$= |E_{od}(x,y)|^2 + {A_r}^2 + E_{od}(x,y)A_r e^{j\phi_r} e^{j(k_{0x}x+k_{0y}y)} + {E_{od}}^*(x,y)A_r e^{-j\phi_r} e^{-j(k_{0x}x+k_{0y}y)}$$

(3-3)

where the proportionality constant has been neglected for convenience. As stated earlier, the recording of the hologram demonstrated above is optically done in experiments, but in simulations Eqs. (3-1)-(3-3) are used for constructing computer-generated holograms (CGHs) numerically [81].

3.1.2 Reconstruction of Intensity Images

The hologram function in Eq. (3-3) has four components. The first two components are the zeroth order information from the intensity of the object and reference; the third component containing the object information is what we desire; the last term is a twin image of the third term. In off-axis holography, the angles where θ_x and θ_y introduce a carrier frequency in the hologram, which can separate the zeroth order, diffracted object field and twin image in spatial frequency domain. By applying a band-pass filter to the Fourier transform of the hologram, the third term $E_{od}(x,y)A_r e^{j\phi_r}$ in Eq. (3-3) can be selected. The filtered spectrum is centered and zero padded to the same size as the original spectrum. For reconstruction, upon multiplying this with the reading (plane) wave, the reconstructed object field at the hologram plane can be obtained. For reconstruction around the object plane, $E_{od}(x,y)$ is back propagated to three planes along the z-axis, a focused plane (red plane in Figure 3-1) and two defocused planes (blue and green planes in Figure 3-1). In conventional reconstruction, the wrapped phase is extracted from the complex reconstructed field at the focused plane and unwrapped using commonly available methods such as PUMA. For DH+TIE, intensity profiles I_n, $n = 0, \pm 1$ are obtained numerically at the focused and defocused planes using the relations

$$I_n(x,y) = I_o(x,y;n\Delta z) \propto |E_{on}{}'(x,y)|^2 \equiv \left|E_{od}(x,y) * h_p(x,y;-d+n\Delta z)\right|^2, n = 0, \pm 1. \tag{3-4}$$

Turning now to in-line holography, a way to find the diffracted object field at the hologram plane is to use the phase-shifting technique [50]. Instead of registering only one hologram, the phase-shifting technique usually requires nominally four in-line holograms with different phase

shifts $0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}$ added to the reference beams. Then, the complex object field at the hologram plane can be obtained from

$$E_{od}(x,y) = \frac{1}{4A_r}\left[\left(h_0(x,y;d) - h_\pi(x,y;d)\right) + j\left(h_{\pi/2}(x,y;d) - h_{-\pi/2}(x,y;d)\right)\right]. \quad (3\text{-}5)$$

As before, the diffracted object field can be back propagated to three planes around the focused plane and the three intensity profiles acquired using Eq. (3-4).

If the phase shifting technique is not used, the second term in the expression for the hologram function in Eq. (3-3) may be possibly eliminated by subtracting the reference intensity, which can also be achieved in experiment by recording the reference intensity at the recording plane. However, this still leaves the first and the last two terms. When back-propagated, and assuming that the object field is much weaker than the reference, the reconstruction results are affected by contributions from the twin image; however, its effect can be minimized if the twin image is severely defocused. The effect of twin image on in-line DH+TIE will be further demonstrated in Section 3.2.

3.1.3 Phase Retrieval Using TIE

TIE is applied to phase retrieval because of its advantages of yielding the unwrapped phase and computational speed. TIE is partial differential equation (for a detailed derivation, see Appendix A)

$$\boldsymbol{\nabla}_\perp \cdot (I\boldsymbol{\nabla}_\perp \phi) = -k_0 \frac{\partial I}{\partial z}, \quad (3\text{-}6)$$

where I denotes the intensity and ϕ is the phase. Under the assumption that the intensity of the object is approximately a constant (as is true for phase objects), Eq. (3-6) can be reduced to a Poisson equation:

$$\nabla^2 \phi = -\frac{k_0}{\bar{I}} \frac{\partial I}{\partial z}, \quad (3\text{-}7)$$

where $\bar{I}$ is the intensity (ideally, constant) at the focus.

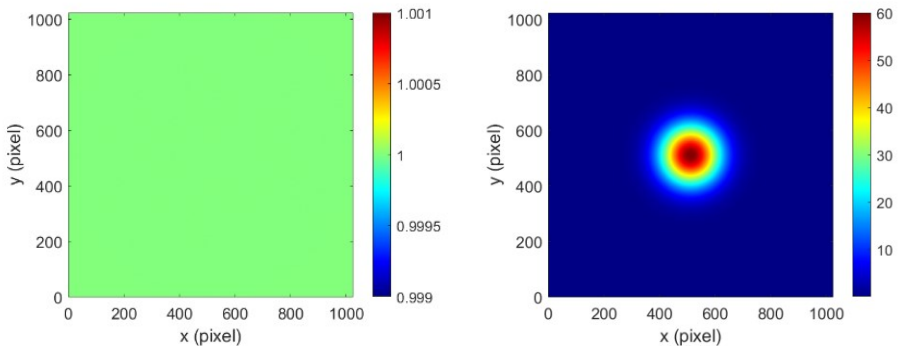
Equation (3-6) can be solved by using the FFT numerically. The longitudinal derivative $\frac{\partial I}{\partial z}$ can be obtained by using finite difference of intensity images along the z-axis. The final retrieved object phase is

$$\phi_o{}'(x,y) = \Im^{-1}\left\{\frac{1}{k_x^2+k_y^2}\Im\left\{\frac{k_0}{\bar{I}}\frac{I_1-I_{-1}}{2\Delta z}\right\}\right\}, \tag{3-8}$$

where $\Im\{\cdot\}$ and $\Im^{-1}\{\cdot\}$ stands for Fourier transform (FT) and inverse Fourier transform (IFT); k_x and k_y are the spatial frequencies along x- and y-axes, respectively. $\bar{I}$ can be a constant, the average value of I_1 and I_{-1}, or I_0. From a preliminary verification, we find that using $\bar{I} = const$ or $\bar{I} = \frac{I_1+I_{-1}}{2}$ has a better accuracy performance than using $\bar{I} = I_0$. Thus, only first two conditions are investigated in the following simulations. Note that if $k_x = 0$ and $k_y = 0$, there will be a singularity. Since this singularity only occurs at the zero spatial frequency, the term $\frac{1}{k_x^2+k_y^2}$ is intentionally set to zero at only the singularity point; this does not have any effect on the evaluation of the IFT integral.

3.2 Simulations and Analysis

In all simulations, unless otherwise stated, the recording distance is set at 200 mm. The pixel pitch is taken to be 4.65 μm. Also, the defocus distance for TIE is taken to be $\Delta z = 0.1$ mm, which is much less than the depth of focus corresponding to the object dimensions studied. The nominal operating wavelength is taken as 514.5 nm. For simplicity, the reference and object intensities are assumed to be unity. This does not violate the usual assumption that the object field should be much weaker than the reference since the magnitude of the *propagated* optical field at the recording plane in our case is much less than unity.



(a) (b)

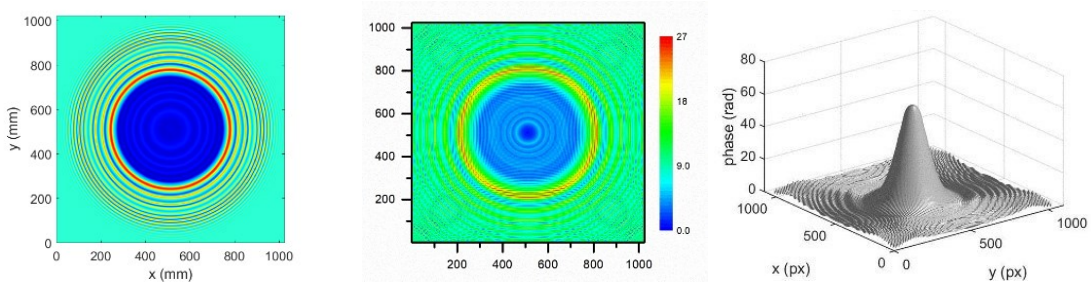
Figure 3-2: The object with a maximum phase value of 60 rad for a wavelength of 514.5 nm and a uniform intensity. (a) The intensity of the object; (b) the phase of the object. [79]

A Gaussian phase object with a uniform (unity) intensity profile is simulated to predict the performance of DH+TIE and compared with SWDH, DWDH results. The size of the object plane is designed to have a resolution of 1024×1024, similar to the resolution of the charged coupled device (CCD) array on the recording plane. The width of Gaussian phase is 0.5 mm (approximately 108 pixels). In Figures 3-2(a,b), the intensity and phase of a typical object are presented with a maximum phase value of 60 rad. The recording distance, i.e., the distance from the object plane to the hologram plane, is set to 200 mm. Then based on Eqs. (3-1)-(3-3), CGHs are computed numerically.

3.2.1 In-line DH+TIE

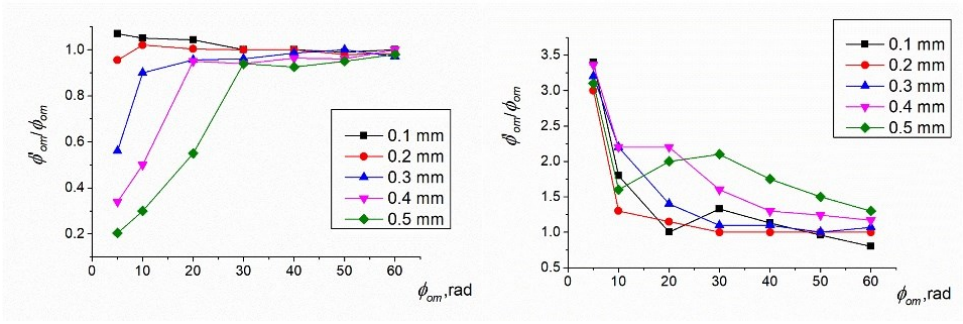
For in-line DH+TIE, reconstruction of the image can be distorted by its twin image and the zeroth order component (see Eq. (3-3), without the second term on the RHS). The effect of the twin image on phase retrieval using DH+TIE for objects with different phase excursions are now studied.

An in-line hologram is depicted in Figure 3-3(a) for the Gaussian object shown in Figure 3-2. The hologram is back propagated to the two planes $z = -\Delta z, \Delta z$. Intensity images at the these planes are shown in Figure 3-3 (b,c). Then, TIE is applied to retrieve the phase.



(a) (b) (c)

Figure 3-3: (a) A CGH generated by an in-line DH of the Gaussian phase object with a maximum phase of 60 rad. (b) Reconstructed intensity profile at $z = \Delta z$ ($z = -\Delta z$ is similar). (c) Retrieved phase using TIE and taking $\bar{I} = 1$. [79]



(a) (b)

Figure 3-4: The figure of merit $\frac{\phi_{om}{}'}{\phi_{om}}$ for Gaussian phase objects with widths $\sigma = 0.1, 0.2, 0.3, 0.4, 0.5$ mm, and peak value $5, 10, 20, 30, 40, 50, 60$ rad, (a) $\bar{I} = 1$, (b) $\bar{I} = \frac{(I_1 + I_{-1})}{2}$. [79]

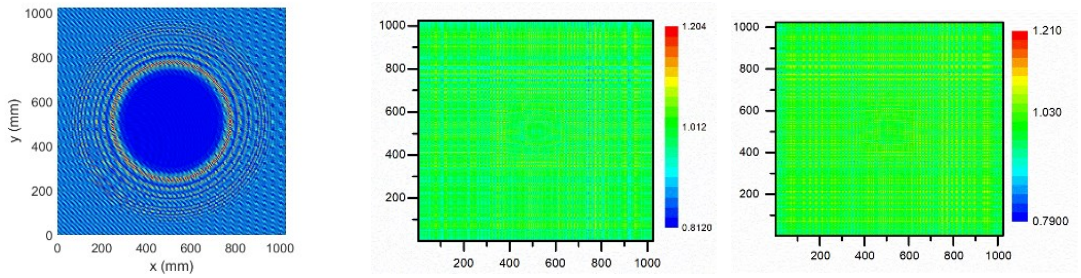
By using Eq. (3-8), the object phase $\phi_o{}'(x, y)$ can be computed from the finite difference of the intensity images. As discussed above, two cases are studied: (a) $\bar{I} = 1$; (b) $\bar{I} = \frac{(I_1 + I_{-1})}{2}$. To investigate the twin image effect on the DH+TIE, Gaussian phase objects with widths $\sigma = 0.1, 0.2, 0.3, 0.4, 0.5\ mm$ and peak values $\phi_{om} = 5, 10, 20, 30, 40, 50, 60$ rad are simulated with in-line DH+TIE. For each of the cases, the reconstructed maximum value $\phi_{om}{}'$ is extracted and compared with the original maximum phase value ϕ_{om}. We define the ratio $\frac{\phi_{om}{}'}{\phi_{om}}$ as a figure of merit to evaluate the reconstruction quality. A ratio close to unity is an indication that the retrieved

phase is close to the original phase. Figures 3-4 (a,b) show the figure of merit for all cases, and when (a) $\bar{I} = 1$, (b) $\bar{I} = \frac{(I_1 + I_{-1})}{2}$.

From Figure 3-4, for both cases, as the peak phase value becomes larger, or the Gaussian phase width is smaller, the figure of merit $\frac{\phi_{om}{}'}{\phi_{om}}$ tends to 1, implying a better reconstruction. From the trends of peak ratios, we can conclude that accuracy of reconstruction with in-line DH+TIE depends on the spread of the phase object and its maximum phase. For widely spread objects and small peak values, the accuracy is low because the influence of the remaining zero-order and, especially, the twin image is very strong. However, for phase objects with large phase gradients, the twin image is severely defocused, thus, the effect on the phase retrieval results is less.

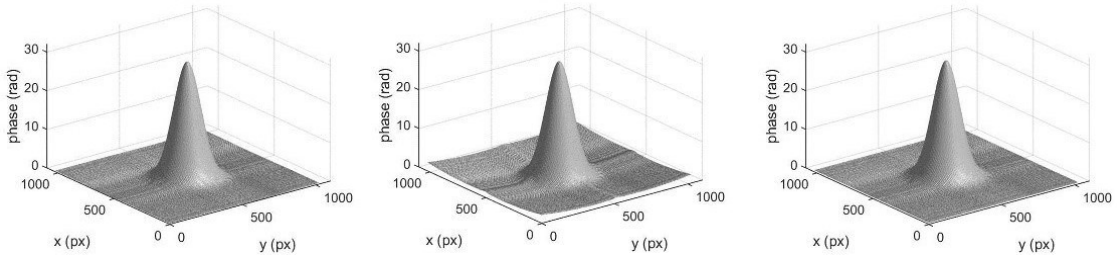
3.2.2 Off-axis DH+TIE

An off-axis hologram is simulated for a Gaussian phase object with $\sigma = 500$ μm (108 pixels) and a peak phase of 30 rad to yield the computer-generated hologram (CGH) as in Figure 3-5(a). An angle of 1.5° both along *x*- and *y*-axis is introduced, thus a carrier frequency, as evidenced by the straight-line fringes, can be observed in the figure. Next, the recorded hologram is Fourier transformed, and a bandpass filter is applied to filter out the zeroth order and twin image in the spatial frequency domain. The resulting filtered hologram is re-centered, zero-padded to the dimensions of the original recorded hologram and reconstructed through back-propagation to the three planes at $z = \Delta z, 0, -\Delta z$. At each of the planes, an intensity image can be calculated by Eq. (3-4). Intensity images at $z = \pm\Delta z$ are shown in Figures 3-5(b,c). Then, the phase can be retrieved by TIE assuming: (a) $\bar{I} = const$; (b) $\bar{I} = \frac{(I_1 + I_{-1})}{2}$. The retrieved (unwrapped) phase profiles ϕ_0' are presented in Figure 3-6(a,b) respectively.



(a) (b) (c)

Figure 3-5: (a) A CGH generated by an off-axis setup DH of the object with uniform intensity and phase profile in (a). The reconstructed intensity images at the planes (b) $z = \Delta z$; (c) $z = -\Delta z$. [79]



(a) (b) (c)

Figure 3-6: Retrieved unwrapped phase by using off-axis DH+TIE (a) $\bar{I} = const$; (b) $\bar{I} = \frac{(I_1 + I_{-1})}{2}$ for a Gaussian phase with maxima radians of 30 rad. (c) Retrieved unwrapped phase using SWDH with PUMA for the same object in (a,b). [79]

For comparison, the reconstructions of object phase using off-axis (single wavelength) SWDH+PUMA (to be called simply SWDH for brevity) are also presented. First, the wrapped phase can be computed from the object field ${E_{oo}}'(x, y)$ at the reconstruction plane $z = 0$ (see Eq. (3-4)) by using:

$$\phi'_{0wrap}(x, y) = \arctan\left(\frac{Im\left({E_{oo}}'(x,y)\right)}{Re\left({E_{oo}}'(x,y)\right)}\right). \tag{3-9}$$

Due to the trigonometric function, ϕ'_{0wrap} is wrapped within the interval $[0,2\pi]$. PUMA [51] is used to unwrap the phase, and the unwrapped phase ϕ'_0 for SWDH is presented in Figure 3-6(c).

To get a better understanding of the performance of phase retrieval, the mean squared error (MSE) defined as

$$MSE = \frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} [\phi_0'(i,j) - \phi_o(i,j)]^2, \tag{3-10}$$

is used as a figure of merit, where M and N are the pixel numbers in x- and y-directions, respectively. In this simulation, $M = N = 1024$. The smaller values of MSE denotes a better reconstruction quality. The MSEs are computed for the two cases of off-axis DH+TIE and off-axis SWDH for a Gaussian phase with $\sigma = 500$ μm, and with peak phase values $5, 30, 60$ rad, and are presented in Table 3-1.

Table 3-1: MSE for off-axis DH and DH+TIE reconstructions

Methods	Mean Squared Error *MSE* (Gaussian phase for $\lambda = 514.5$ nm)		
	(5 rad)	(30 rad)	(60 rad)
Off-axis DH+TIE ($\bar{I} = const$)	0.0098	0.0070	0.0039
Off-axis DH+TIE $\left(\bar{I} = \frac{I_1 + I_{-1}}{2}\right)$	0.0697	0.0843	0.1118
Off-axis SWDH	0.0069	0.0070	0.0070

From Table 3-1, the MSEs indicate off-axis DH+TIE and SWDH have comparable accuracy performance. For off-axis DH+TIE, the assumption that $\bar{I}$ is a constant can lead to a better result and the accuracy increases as the peak value of the Gaussian phase becomes larger.

As is well-known, the use of dual-wavelength (λ_1, λ_2) DH enables minimizing or even eliminating the need for phase unwrapping due to a larger synthetic wavelength. We have therefore also compared our off-axis DH+TIE results with the results from DWDH+PUMA (again only called DWDH for brevity) and SWDH. For DWDH, the wavelengths $\lambda_1 = 514.5$ nm and $\lambda_2 = 496.5$ nm have been used. The wrapped phase map of two wavelengths are subtracted to obtain a phase map corresponding to the synthetic wavelength $\Lambda = \frac{\lambda_1 \lambda_2}{|\lambda_1 - \lambda_2|} = 14.19$ μm [6]. The difference

in wavelength introduces a scaling effect to the holograms, and this is compensated by zero padding the hologram for one wavelength and the wrapped phase for the other wavelength. The pad size is given by [6]

$$\text{pad size} = \text{round}\left(\frac{N}{2}\left(\frac{\lambda_1}{\lambda_2} - 1\right)\right). \tag{3-11}$$

Simulations done for a Gaussian phase with a peak of 30 rad and width of $\sigma = 500$ μm, not shown here, show that the results of DH+TIE with $\bar{I} = const$ and SWDH match the best, while that for DWDH suffer from higher mean square errors due to (i) peak to peak retrieved phase not being exactly the same as in the other two cases, and (ii) some random fluctuations in the retrieved phase. It is surmised that these random variations in simulations may arise from round-off errors stemming from Eq. (3-11); this is still under investigation. Comparison between DH+TIE, SWDH and DWDH appears later in connection with actual experiments performed on phase retrieval.

3.2.3 PSDH+TIE

To complete the list of existing methods for phase retrieval, we consider PSDH and introduce the concept of combining PSDH with TIE. As mentioned earlier, PSDH is an in-line holographic technique where the reference with nominally four different phase shifts, viz., $0, \frac{\pi}{2}, \pi, -\frac{\pi}{2}$ are used to generate four holograms. By using Eq. (3-5), the complex object field at the hologram plane is obtained. The complex field is back propagated to two planes at $z = -\Delta z, \Delta z$, as shown in Figures 3-7(a,b). Then, TIE is applied to the intensity images for (1) $\bar{I} = const$; (2) $\bar{I} = \frac{(I_1 + I_{-1})}{2}$. The reconstructed results are shown in Figures 3-7(c,d) respectively. For comparison, the traditional PSDH+PUMA (abbreviated as PSDH) is also performed for the object and the retrieved unwrapped phase using PUMA is shown in Figure 3-7(e).

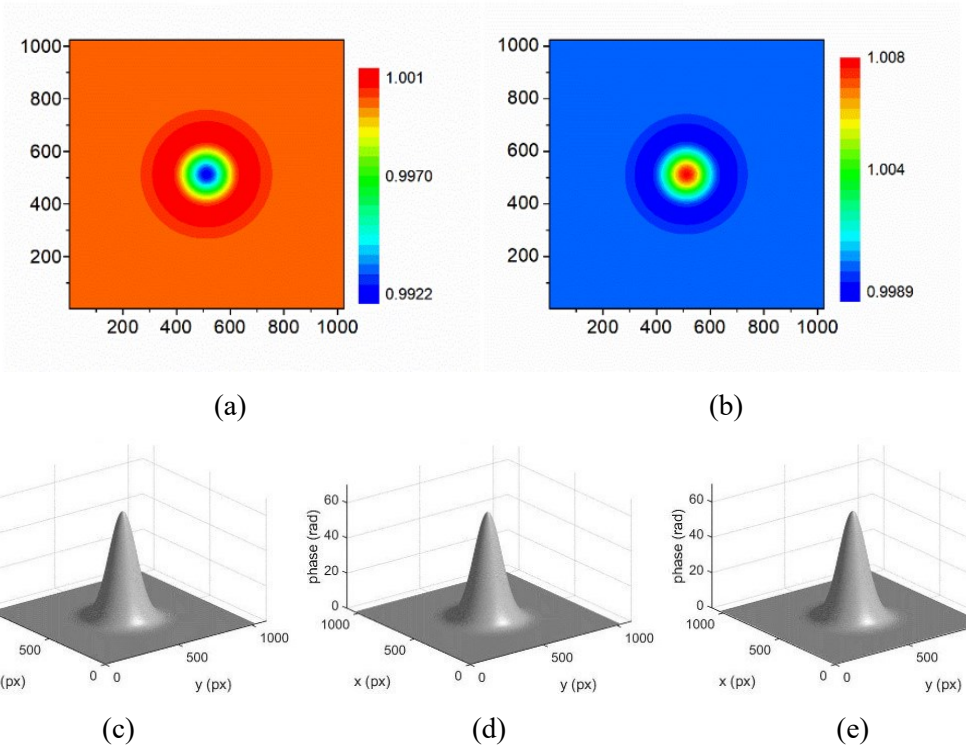


Figure 3-7: The reconstructed intensity profiles at (a) $z = -\Delta z$; (b) $z = \Delta z$. Retrieved phase by using PSDH+TIE under the condition (c) $\bar{I} = const$; (d) $\bar{I} = \frac{(I_1 + I_{-1})}{2}$. (e) Retrieved phase by using PSDH with traditional phase unwrapping using PUMA. [79]

Table 3-2: MSE for PSDH+TIE and PSDH phase retrievals

Methods	Mean Squared Error *MSE* (Gaussian phase for $\lambda = 514.5$ nm)		
	(5 rad)	(30 rad)	(60 rad)
PSDH+TIE	3.63×10^{-10}	4.04×10^{-9}	2.25×10^{-8}
PSDH	4.63×10^{-5}	4.64×10^{-5}	4.65×10^{-5}

Since both conditions, viz., $\bar{I} = const$ and $\bar{I} = \frac{(I_1 + I_{-1})}{2}$ in Figures 3-7(c,d), respectively, yield results to the same level of accuracy, we choose the condition $\bar{I} = const$ to compare with PSDH results. MSEs are calculated for the simulations of Gaussian phase objects with peak phase of $30, 45, 60$ rad and width $\sigma = 500$ μm. The calculated MSEs are listed in Table 3-2. The MSE values in Table 3-2 are extremely small, which shows both PSDH+TIE and PSDH can give high

quality phase retrieval results theoretically; however, PSDH+TIE is more accurate compared to PSDH.

3.2.4 Computational Performance

It is in the computation time that phase retrieval using TIE shows its definite advantage. A time performance evaluation is done for SWDH (+PUMA) and DH+TIE. The numerical phase retrieval and reconstruction is performed by Matlab R2019b using a computer with Intel(R) Core (TM) i7-9700 CPU @ 3.00 GHz. The computation time is averaged over 5 runs. As shown in Figure 3-8, Gaussian phase objects with peak phase varying from 5 rad to 60 rad are retrieved using off-axis and in-line SWDH and DH+TIE. Compared to off-axis SWDH, in-line SWDH requires more time, which may possibly result from PUMA requiring additional computation time due to distortions related to the existence of the twin images. DH+TIE is far superior to DH+PUMA methods in time performance. Particularly, as the peak phase becomes larger, DH+PUMA methods require substantially more computation time; in contrast, DH+TIE still shows very good time efficiency, and computational times are independent of the peak phase excursion.

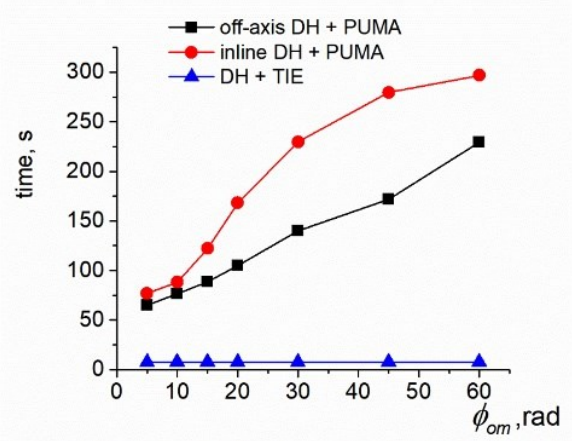


Figure 3-8: Time performance for phase retrieval using in-line and off-axis SWDH+PUMA and DH+TIE from CGHs for Gaussian phase object with a peak phase varying from 5 rad to 60 rad. [79]

3.3 Experimental Results

3.3.1 Experimental Setup

A typical off-axis setup for holographic recording in transmission geometry is used to digitally record holograms of phase objects as shown in Figure 3-9 [6]. Illumination is provided by an Ar-ion laser at 514.5 nm for SWDH, while an additional wavelength of 496.5 nm is used for DWDH. The laser beam is collimated and expanded by a spatial filter assembly and a collimation lens with a focal length of 250 mm. Then a Mach-Zehnder interferometer setup is used to obtain the interference pattern between the object beam and the reference beam. By adjusting the aperture size of the two irises, the interference pattern can cover the whole active area of the CCD camera.

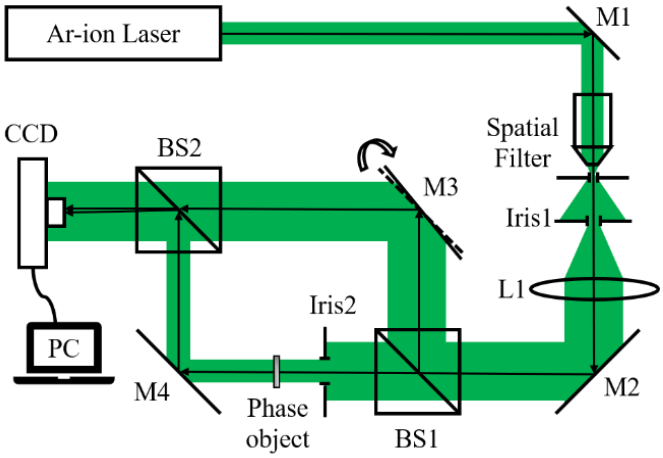


Figure 3-9: A schematic diagram of the optical system used in the experiments. An Ar-ion laser with wavelengths of $\lambda_1 = 514.5$ nm and $\lambda_2 = 496.5$ nm is used. A spatial filter assembly, composed of a $10\times$ microscope objective and a 25 μm pinhole, together with a collimation lens $L1$ is used to collimate and expand the laser beam. A Mach-Zehnder interferometric setup is used. The reference beam path follows BS1-M3-BS2, while the object beam path is BS1-M4-BS2. A small angle between the reference and object beams can be introduced by slightly tilting the mirror $M3$ to achieve the off-axis setup. Two irises are used to control the size of the laser beam. L: lens; M: mirror; BS: beam splitter; CCD: Charged Coupled Device; PC: personal computer. [79]

The distance from the object to the CCD camera is set to 200 mm. The CCD camera (Thorlabs, DCU223C) has a resolution of 768×768 pixels and a pixel pitch of 4.65 μm in both x- and y-axis, which corresponds to an area of 3.57×3.57 mm^2. A small angle is introduced to the reference

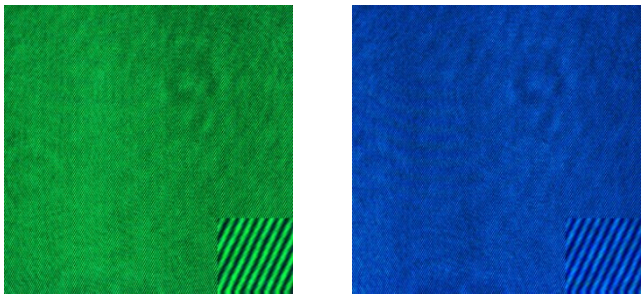
beam to spatially discriminate reconstructed fields and zeroth term. The angle is usually within 3°; the choice of the angle depends on the wavelength and pixel pitch of the CCD camera [83].

3.3.2 Phase Retrieval Results and Analysis

A set of plano-convex lenses are used as the phase objects in the experiments, in order to investigate the performance of DH+TIE, SWDH and DWDH for different phase excursions.

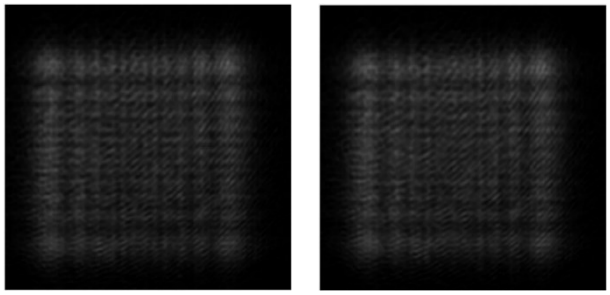
Figures 3-10(a,b) are two representative holograms recorded by the CCD camera with wavelengths of 514.5 nm and 496.5 nm respectively. The phase object used is a lens with a radius of curvature of $R_1 = 516.8$ mm, a refractive index of $n = 1.52$, and an aperture diameter of $D = 25.4$ mm. To mitigate the effect from the aberration of the lens, only a small part of the central part of the lens has been illuminated. The diffracted field covers the whole area of the CCD camera.

A similar reconstruction and phase retrieval process as discussed in Section 3.1 and 3.2 is also done for the experimental data. After processing in spatial frequency domain and back propagating to two planes close to the reconstruction plane, two intensity images are computed numerically at $z = -\Delta z,\ \Delta z$, respectively (Figures 3-11(a,b)). As noted earlier, the value of $\Delta z = 0.1$ mm is used.



(a) (b)

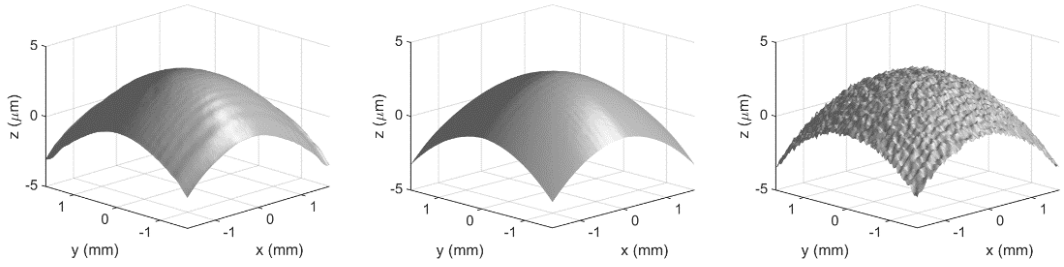
Figure 3-10: The holograms of recorded by the CCD camera with a wavelength of (a) $\lambda_1 = 514.5$ nm; (b) $\lambda_2 = 496.5$ nm. The inset is a zoom in view for the hologram showing the carrier frequency by the off-axis setup. [79]



(a) (b)

Figure 3-11: The two reconstructed intensity images at $z =$ (a) $-\Delta z$; (b) $+\Delta z$. [79]

The obtained intensity images can be used in TIE to retrieve the phase for the lens. Given that a constant intensity for the phase object is not achievable in experiment, $\bar{I} = \frac{I_1 + I_{-1}}{2}$ is used in TIE. Also, the phase is converted to height with the known value of refractive index and the corresponding wavelength. Figure 3-12(a) shows the retrieved profile of the lens by using DH+TIE followed by phase to height conversion. From the hologram in Figure 3-10(a), SWDH can be applied to reconstruct the lens. The profile from SWDH is shown in Figure 3-12(b). Also, reconstructed from holograms in Figures 3-10(a,b) respectively, two phase maps for two different wavelengths of 514.5 nm and 496.5 nm are derived and subtracted to yield the phase map corresponding to the synthetic wavelength $\Lambda = 14.19$ μm. The retrieved profile from DWDH is shown in Figure 3-12(c).

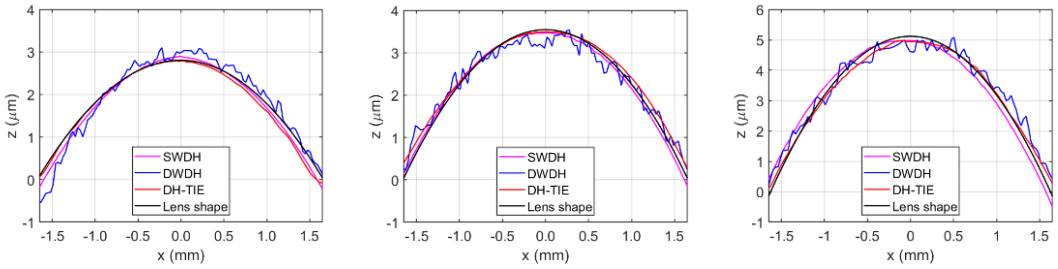


(a) (b) (c)

Figure 3-12: Three-dimensional topography of plano-convex lens with $R_1 = 516.8$ mm by using (a) DH+TIE; (b) SWDH; (c) DWDH. In (b) and (c), PUMA is used for phase unwrapping. [79]

For comparison, a line along the x-axis and passing through the center of the lens is extracted for each of the methods in Figure 3-12. Figure 3-13(a) shows the extracted profile along x for $y = 0$ and the ideal lens profile. From this plot, it is seen that all the methods perform very well, though the result from DWDH is noisier than the other two methods. This is in general agreement with our results from CGH reconstructions in Section 3.2. Thereafter, two other lenses with same aperture size and refractive index but smaller radii of curvatures are used as phase objects. The radii of curvature for the two lenses are $R_2 = 387.6$ mm and $R_3 = 258.6$ mm, respectively. The difference in radii of curvatures lead to a change in the peak phase (and the phase gradients) in the transverse plane. The same experimental recording and numerical reconstruction process is also performed for these two lenses. Figures 3-13(b,c) present the extracted lines for each of these lens by using all the three methods and the ideal lens profile.

The figure of merit MSE given in Eq. (3-10) is calculated for each of the methods and are shown in Table 3-3. MSEs give an insight on the quality of phase retrieval. From the MSEs, DH+TIE and SWDH give comparable quality of results. DWDH is noisier than the other two methods. The experimental results are in accordance with our simulations in Section 3.2.



(a) (b) (c)

Figure 3-13: Extracted lines along x-axis and go through the center of the lens for lens with a radius curvature of (a) $R_1 = 516.8$ mm; (b) $R_2 = 387.6$ mm; (b) $R_3 = 258.6$ mm. The magenta line is the result from SWDH; the blue line is extracted from DWDH result; the red line is the result from DH+TIE; and the black line is the ideal lens shape. For SWDH and DWDH, PUMA is used for phase unwrapping. [79]

Table 3-3: MSEs in heights for experimental results

Methods	Mean Squared Error *MSE* for heights [μm^2]		
	$R = 516.8$ mm	$R = 387.4$ mm	$R = 258.6$ mm
DH+TIE	0.0463	0.0997	0.1352
SWDH	0.0710	0.0880	0.2832
DWDH	0.1062	0.1244	0.5494

As mentioned in Section 3.2, phase retrieval results from DWDH in simulations using CGHs were noisier than SWDH and DH+TIE methods. A probable reason was attributed to inaccuracies from zero-padding. In actual experiments, the noise may also arise from the laser linewidth. In SWDH, the fractional error in measurement of phase is of the order of $\frac{|\delta\varphi_\lambda|}{\varphi_\lambda} = \frac{|\delta\lambda|}{\lambda}$, where $|\delta\lambda|$ is the laser linewidth. In DWDH, the fractional error in measurement of phase is $\frac{|\delta\varphi_\Lambda|}{\varphi_\Lambda} = \frac{|\delta\Lambda|}{\Lambda}$. Using the expression for the synthetic wavelength mentioned in Section 3.2, viz., $\Lambda = \frac{\lambda_1\lambda_2}{|\lambda_1-\lambda_2|} \approx \frac{\lambda^2}{\Delta\lambda}$, it can be shown that $\frac{|\delta\varphi_\Lambda|}{\varphi_\Lambda} \approx 2\frac{|\delta\lambda|}{\lambda} + \frac{|\delta\lambda|}{\Delta\lambda} \approx \frac{|\delta\lambda|}{\Delta\lambda}$. Therefore, the ratio of the errors is $\frac{Error_{DWDH}}{Error_{SWDH}} \approx \frac{\lambda}{\Delta\lambda}$. In our experiment, this ratio is of the order of 25, indicating that DWDH is expected to have much more error than SWDH.

3.3.3 Computational Performance

In Section 3.2.4, a computational performance analysis was done to compare phase retrieval from CGHs using DH+TIE with other SWDH methods using PUMA. A similar computation time comparison is done for the experimental results. Once again, numerical phase retrieval and reconstruction is performed by Matlab R2019b using a computer with Intel(R) Core(TM) i7-9700 CPU @ 3.00 GHz. The computation time is averaged over 5 runs. Expectedly, DH+TIE shows its great advantage in time performance with a computational time ~0.5 s, as is presented in Table 3-4. DWDH+PUMA requires about 20 s for reconstruction for all three lenses (phases). The

computational time for SWDH+PUMA is in the order of a few minutes and increases with the peak phase (reduction of radius of curvature of lens), since more unwrapping is needed.

Table 3-4: Computational time for phase retrieval from experimental results

Methods	Computation Time [s]		
	$R = 516.8$ mm	$R = 387.4$ mm	$R = 258.6$ mm
DH+TIE	0.498	0.499	0.499
SWDH	124.103	132.983	164.398
DWDH	20.902	20.032	21.202

3.4 Conclusion

In this Chapter, the general idea of DH+TIE is introduced, and a detailed description of the method is described. Its performance is evaluated using convenient figures of merit such as MSE and computational time. Specifically, conventional DH techniques such as SWDH and DWDH using PUMA for phase unwrapping have been compared with in-line and off-axis DH+TIE, along with a newly introduced PSDH+TIE technique. Both CGHs as well as experimentally recorded holograms have been used for phase retrieval. The advantages of using TIE coupled with DH as a convenient tool for phase retrieval are that multiple registrations of images are automatically done numerically with a single hologram capture and without the need for moving optical components, the computed phase is naturally unwrapped, and computation times are much smaller than PUMA.

A few of the other results of the work are now listed. The twin image effect on in-line DH+TIE has been studied. For objects with larger phase excursions, the twin image is severely defocused; thus, in-line DH+TIE can perform better for large phase radians cases. Off-axis DH+TIE is compared with standard DH methods in simulation and experiments. Though SWDH has an accuracy performance comparable with DH+TIE, the computation time for SWDH is much higher. DWDH can potentially avoid/minimize phase unwrapping iterations; however, its accuracy is not

as good as DH+TIE – possible reasons being round-off errors during zero padding and the use of dual wavelengths. The idea of combining PSDH with TIE is also proposed in this Chapter. Our simulations with CGHs indicate PSDH+TIE can achieve excellent accuracy in an in-line geometry, albeit with multiple captures of the holograms.

In the future, DH+TIE can be further developed by using higher-order intensity derivatives in TIE [80], non-recursive and recursive methods for combining TIE with TPE (see Chapter II and Ref. [36]), etc. It is also worthwhile to explore DH+TIE using partially coherent light with the help of optical scanning holography [84] or Fresnel incoherent correlation holography (FINCH) [85]. The next chapter will discuss the application of correlation of holograms in 3D surface feature identification.

CHAPTER IV

DIGITAL CORRELATION OF CGHS FOR 3D FACE RECOGNITION

In this Chapter, we explore a potentially important application of DH in identification of 3D features from hologram correlation without the need for any reconstruction. This is possible since the 3D information of objects is encoded in the 2D hologram. Therefore, digital hologram correlation can reveal critical 3D surface features by comparing with a reference correlation result. This technique, first developed to compare surface roughness during 3D metallic additive manufacturing [82], is applied to facial recognition.

4.1 Methodology

4.1.1 Pre-processing of 3D Face Models

The original human face point clouds used for simulations were acquired by a multi-view camera with structured light imaging, kindly provided by Prof. Yebin Liu' group at Tsinghua University. Each face point cloud roughly contains more than 100,000 points, covering the whole face part (eyes, nose, mouth), and some parts of the hair, ears, and shoulders. In order to uniformly compare all the face models, pre-processing is manually done to align the faces and facing directions, crop redundant parts of the model, mesh the point cloud with triangles in Meshlab® and extract the feature points for each human face. The feature points are the feature map of a human face. A typical pre-processed image is shown in Figure 4-1. The 32 red spheres represent 32 feature points. Points 1-10 show the contour of the face; points 11-22 gives the perimeter of the two eyes; points 23-26 represent the nose bridge; and points 27-32 are six points on the contour of the mouth. These feature points contain typical 3D depth and intensity.

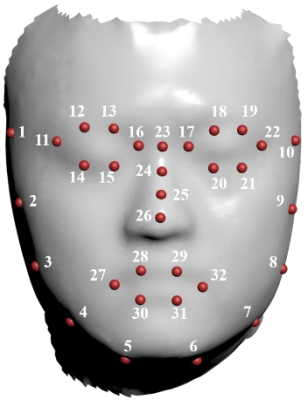


Figure 4-1: A schematic diagram of a pre-processed face model. The red spheres labeled from 1 to 32 represent the feature points of the face. Points 1-10 is the contour of the face; points 11-22 gives the perimeter of the two eyes; points 23-26 represents the nose bridge; and points 27-32 are six points on the contour of the mouth. [81]

4.1.2 Calculations of the CGHs

Based on selected feature points, three CGH algorithms are implemented respectively. One way to generate CGHs is a point-oriented algorithm [86-88] to treat every feature point as a point source and propagate these points (considered as point sources) by Rayleigh-Sommerfeld diffraction. Another way is an angular spectrum layer-oriented algorithm [89], which divides a 3D complex scene into multiple layers according to depth cues. The third way is also a point-oriented algorithm based on the Fresnel approximation [90]. The Rayleigh-Sommerfeld point-oriented algorithm has the advantage in accurately interpreting point source propagation. The angular spectrum layer-oriented algorithm shows great superiority in computing efficiency, which is above 1000 times faster than the point-oriented models [89]. Fresnel approximation is a compromise between the first two algorithms with regard to computational efficiency.

The feature points on the face model are considered as a collection of point sources, shown in Figure 4-2, whose holograms can be found by the point oriented or layer-oriented algorithms.

For the point-oriented approach, each point propagates as a spherical wave and is recorded at the hologram plane. The yellow star in the rectangular box represents point (source) 22 which

propagates along the z direction and its CGH is computed at the hologram plane. Then, a wavefront superposition of all propagated point sources is calculated and yields the complex amplitude of the CGH for a corresponding face. The phase hologram can be obtained by:

$$\phi(x,y) = arctan\left(\frac{Im(h(x,y))}{Re(h(x,y))}\right), \tag{4-1}$$

where $h(x,y)$ is the complex hologram (similar to Eq. (3-5)), and $\phi(x,y)$ is a presentation of the phase hologram.

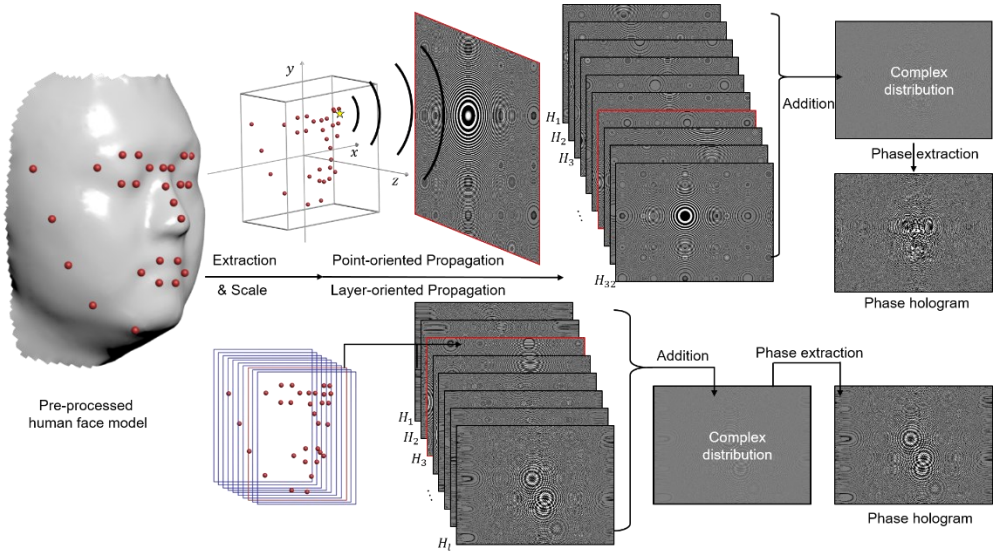


Figure 4-2: A diagram of point-oriented and layer-oriented method for calculating phase holograms. [81]

Figure 4-3 shows the coordinate system for point-oriented algorithms. A collection of feature points is considered as a group of self-illuminated point sources, emitting a series of spherical waves. These spherical waves are interfered and recorded at the hologram plane. In Figure 4-3, $P_i(x_i, y_i, z_i)$ is the i^{th} point near object plane; $P(x, y, d)$ stands for arbitrary point in the hologram plane; the distance between P and P_i is marked as r_i.

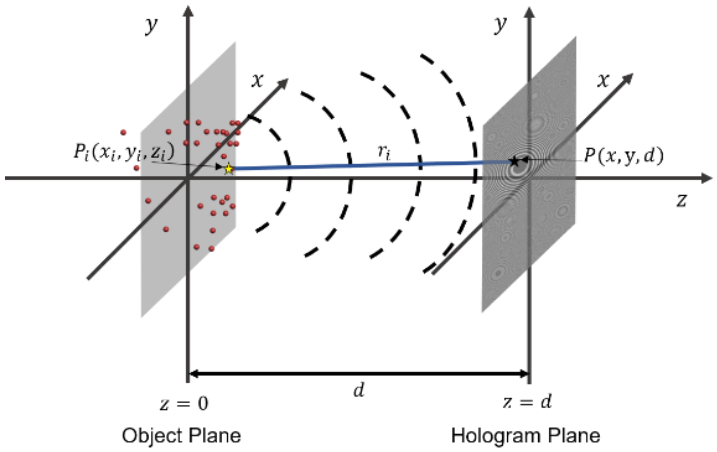


Figure 4-3 Schematic diagram of a point-oriented propagation coordinates. The black dashed lines are the wavefront of the point source (star). [81]

The propagation of Rayleigh-Sommerfeld diffraction is [27]:

$$h(x,y) = j\frac{k_0}{2\pi}\sum_{i=1}^{32}\frac{A_i}{r_i}\exp\left[-j(k_0 r_i + \phi_i)\right], \tag{4-2}$$

where $h(x,y)$ is the complex amplitude distribution of the CGH, A_i is the initial amplitude of the i^{th} point, $k_0 = \frac{2\pi}{\lambda}$ is the wave number, λ is the wavelength of the light, ϕ_i is the initial phase (ϕ_i is set to a constant value in this case), r_i is the distance between the i^{th} point $P_i(x_i, y_i, z_i)$ and an arbitrary point $P(x, y, d)$ in the hologram plane where d is the distance between object plane and hologram plane:

$$r_i = \sqrt{(x_i - x)^2 + (y_i - y)^2 + (d - z_i)^2}. \tag{4-3}$$

In the layer-oriented approach, as shown in Figure 4-2, a set of feature points are sliced into l layers, according to the depth values (z axis). Each layer contains a few point sources that have similar depth values. The complex amplitude distribution is calculated by the angular spectrum method:

$$h(x,y) = \sum_{i=1}^{l}\mathcal{F}^{-1}\{H_{pi}\cdot\mathcal{F}[A_i(x,y)\exp(-j\phi_i(x,y))]\}, \tag{4-4}$$

where $\mathcal{F}$ and $\mathcal{F}^{-1}$ are Fourier and inverse Fourier transform respectively, $A_i(x, y)$ represents the amplitude distribution function of the initial field from the point sources in the i^{th} layer, ϕ_i represents the initial phase distribution, l is the number of layers and H_{pi} is the angular spectrum transfer function for propagation:

$$H_p(k_x, k_y; \Delta z) = \exp\left[-j\left(\sqrt{{k_0}^2 - ({k_x}^2 + {k_y}^2)}(d - z_i)\right)\right]. \tag{4-5}$$

Figure 4-4 shows typical face models with extracted feature points and corresponding CGHs generated by the angular spectrum layer-oriented method [89]. Face 1 and Face 6 are obtained from the same person but with distinctive facial expressions; the other three faces shown in Figure 4-4 are acquired from different individuals.

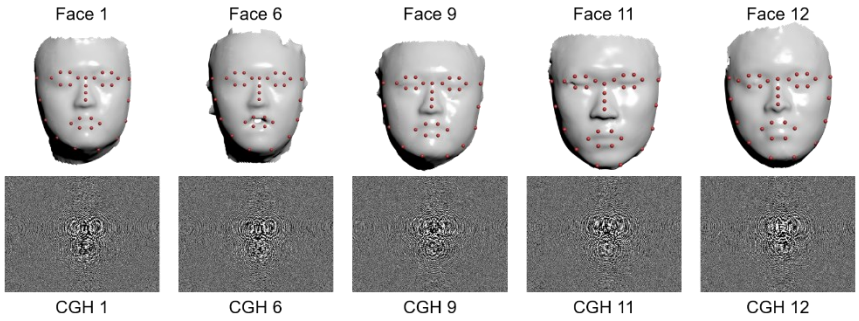


Figure 4-4: Five typical face models and corresponding holograms. Face 1 and Face 6 are the faces from the same person with different facial expressions. Face 9, 11, 12 are the faces from different individuals. [81]

For phase-only hologram, the initial phase is usually taken as random [89,91-92], in order to reduce the low frequency part of the CGH contributed by the modulation of amplitude distribution. We have taken the initial phase as a constant because in our case the number of points for face recognition is small. The propagation and interference of these points cannot contribute a large variance in amplitude distribution of a CGH. The randomness could introduce difficulty in correlation recognition and should be avoided.

Finally, the Fresnel diffraction method is achieved by approximation of r_i in Rayleigh-Sommerfeld method (see Eq. (4-2)), through applying the paraxial approximation and retaining the first two terms of the formula:

$$r_i \approx (d - z_i)\left[1 + \frac{1}{2}\left(\frac{x_i - x}{d - z_i}\right)^2 + \frac{1}{2}\left(\frac{y_i - y}{d - z_i}\right)^2\right]. \quad (4\text{-}6)$$

Then, Eq. (4-2) can be expressed as [27]:

$$h(x,y) = j\frac{k_0}{2\pi}\sum_{i=1}^{32}\frac{A_i \cdot exp\,[-jk_0(d-z_i)]}{(d-z_i)}\,exp\,\left[-j\left(\frac{k_0}{2(d-z_i)}((x_i - x)^2 + (y_i - y)^2) + \phi_i\right)\right]. \quad (4\text{-}7)$$

According to Eq. (4-7), Fresnel diffraction can be used in a point-oriented scenario. Note that scaling and rotating of 3D faces will affect the results. This is because the Fresnel diffraction formula can be recast in terms of a scaled Fourier transform [49]. Since Fourier transformation is dependent on scaling and rotation, the performance of the proposed method will be affected. This is also the reason why we align all the faces in pre-processing.

4.1.3 Correlation and Performance Metrics

Since 3D information of human faces is encoded in 2D CGHs, it is postulated that recognition can be achieved by 2D digital correlation of CGHs. The correlation concept was, in fact, also used for estimating the similarity of other 3D objects [83]. Correlation peak values are extracted as a metric for evaluating similarity of human faces.

Given that all the CGH algorithms have their own strengths, drawbacks and their own propagation range, we have tried and compared them in the application of 3D face recognition in this work. Correlation performance metrics, known as figures of merits [93,94], viz. discrimination ratio (DR), peak-to-correlation plane energy (PCE), and peak-to-noise ratio (PNR) are studied for evaluating the quality of the correlation [95-99]. The DR is the ratio of the autocorrelation peak value to the cross-correlation peak value, which means the ratio between desired target peak value and nontarget peak value. The PCE is defined as correlation peak value over correlation plane

energy, which denotes the amount of energy accumulated in the correlation peak. PCE can also measure the sharpness of the correlation peak [100]. The PNR is the correlation peak value over the mean energy of the correlation plane, showing the correlation peak value compared to the mean noise of the correlation plane. Correlation coefficients $C(m,n)$ is defined in Eq. (4-8), where X_1 and X_2 are the two matrix, k, l are the discrete index along *x,y*-directions, the ranges of m, n are to make the correlation result matrix size same as the original matrix size. Table 4-1 presents a summary for figures of merits used in this work, where P_{Auto} is the autocorrelation peak value, P_{Cross} is the cross-correlation peak value, $C(m,n)$ is the correlation coefficient, M and N are number of pixels for CGH in x and y coordinates, respectively.

$$C(m,n) = \left|\sum_{k=0}^{M-1}\sum_{l=0}^{N-1} X_1(k,l)X_2(k-m,l-n)\right|, \begin{cases} -\frac{M}{2} \le m \le \frac{M}{2}-1 \\ -\frac{N}{2} \le l \le \frac{N}{2}-1 \end{cases}. \quad (4\text{-}8)$$

Table 4-1: A summary of performance metrics for correlation

Metrics	Definitions
Discrimination ration (DR)	$\mathrm{DR} = \frac{P_{Auto}}{P_{Cross}}$
Peak-to-correlation plane energy (PCE)	$\mathrm{PCE} = \frac{P_{Auto}}{E} = \frac{P_{Auto}}{\sum_{m=1}^{M}\sum_{n=1}^{N}\lvert C(m,n)\rvert^2}$
Peak-to-noise ratio (PNR)	$\mathrm{PNR} = \frac{P_{Auto}}{\bar{E}} = \frac{P_{Auto}}{\frac{1}{M\cdot N}\sum_{m=1}^{M}\sum_{n=1}^{N}\lvert C(m,n)\rvert^2}$

4.2 Results and Analysis

The size of the hologram plane is set as 640×480, with a pixel pitch of $\Delta x = \Delta y = 8 \times 10^{-3}$ mm. In order to satisfy the sampling theorem in numerical diffraction and propagation [101], the 3D face feature points in *x-y* plane is scaled to a similar size as the hologram plane. Along z, the propagation distance from $z = 0$ to hologram plane $z = d$ varies among different CGH algorithms.

For the Rayleigh-Sommerfeld point-oriented algorithm, there is no restriction of the propagation distance. We set $d = 500$ mm. For angular spectrum layer-oriented method, due to the Fast Fourier Transform algorithm (FFT) and Nyquist sampling theorem, d has limitations [89]:

$$d \leq \max\left\{M\Delta x\sqrt{4(\Delta x)^2\lambda^{-2}-1}, N\Delta y\sqrt{4(\Delta y)^2\lambda^{-2}-1}\right\}. \tag{4-9}$$

In our case, with $\lambda = 532 \times 10^{-6}$ mm, thus, $d \leq 153.9$ mm. To meet the requirements of the propagation distance, we set $d = 150$ mm. Finally, the limitation of d for Fresnel propagation comes from the approximation in Eq. (4-6). The third term of the paraxial approximation is supposed to be much smaller than $\frac{1}{k_0 z}$, which yields:

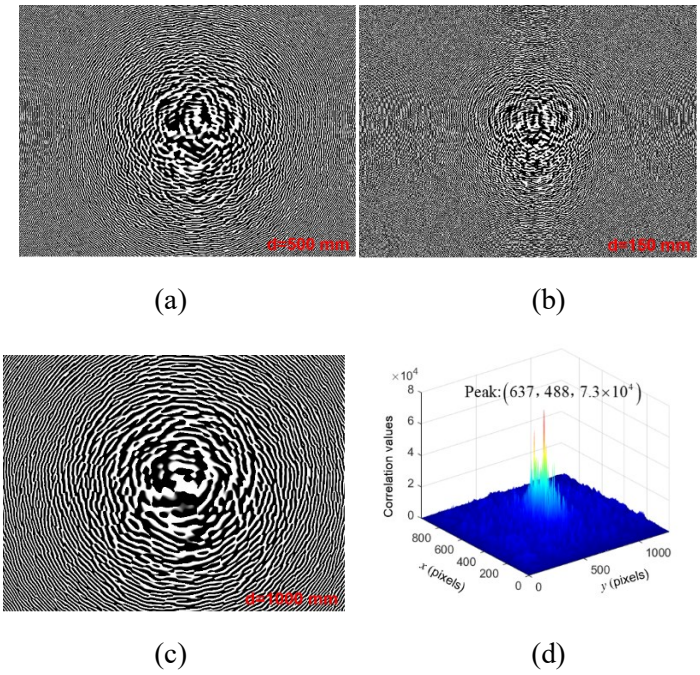
$$d^3 \gg \frac{\pi}{4\lambda}(L_0^2 + W_0^2 + L^2 + W^2)^2, \tag{4-10}$$

where $L = M\Delta x$, $W = N\Delta y$, L_0 and W_0 are length and width of feature points in *x*-*y* plane, which have similar values as L and W. Therefore, $d \gg 200$ mm, and $d = 1000$ mm is used in Fresnel diffraction case.

4.2.1 Face Verification and Identification

Twenty sets of feature points for 20 faces labeled from 1 to 20 are investigated. Face 1 and Face 6 are obtained from the same person but with different facial expressions; other 18 faces are acquired from the other 18 different persons.

In face verification, Face 1 is set as a reference and its CGHs computed by using three CGH algorithms with corresponding propagation distances are shown in Figure 4-5(a-c). The reference CGH is then correlated with the other CGHs. A typical cross-correlation pattern is shown in Figure 4-5(d). Then, the correlation peak value is extracted for further evaluations.



(a) (b)

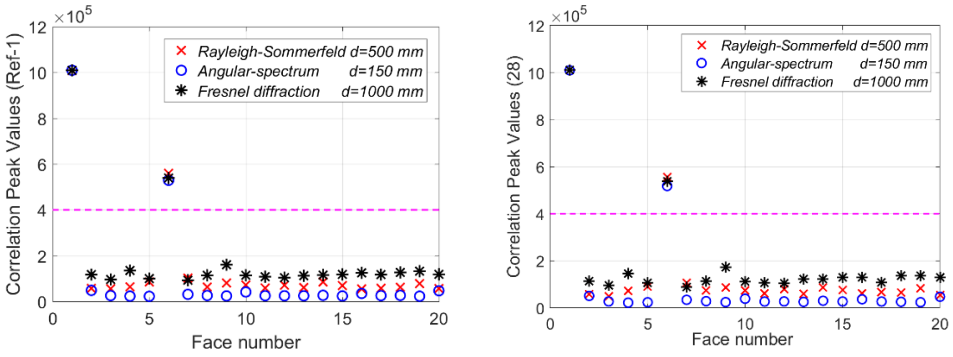
(c) (d)

Figure 4-5: (a) A CGH generated by Rayleigh-Sommerfeld point-oriented algorithm with $d = 500$ mm; (b) A CGH generated by angular spectrum layer-oriented algorithm with $d = 150$ mm; (c) A CGH generated by Fresnel diffraction point-oriented algorithm with $d = 1000$ mm; (d) a typical cross-correlation result for two CGHs. The correlation peak value is 7.3×10^4. [81]

Using each of the three methods for generating CGHs discussed in Section 4-1, the CGH of Face 1 is used as a reference CGH and is sequentially correlated with other CGHs one-by-one. The correlation peak values for each 1:1 correlation are extracted from the correlation patterns. The extracted correlation peak values are plotted versus the face number. In Figure 4-6(a), the correlation values for Rayleigh-Sommerfeld point-oriented algorithm with $d = 500$ mm are labeled in red crosses; the peak values for angular spectrum layer-oriented algorithm with $d = 150$ mm are shown in blue circles; the black asterisks denote the result for Fresnel diffraction point-oriented algorithm with $d = 1000$ mm. The autocorrelation peak value for the reference CGH is in the order of 10^6, while cross-correlation peak values are generally much lower. A noticeable standout is the peak value of correlation between Face 1 and Face 6, which is much higher than the other correlation peak values. As mentioned previously, Face 1 and Face 6 are derived from the same person but with different facial expressions. The correlation peak value of

Face 1 and Face 6 indicates that the two faces share considerable similarity in 3D space. One may manually set a threshold, e.g., the dashed magenta line drawn in Figure 4-6(a), for verification of the reference face. If the correlation peak value is larger than the threshold, the target face can be recognized as the same face as the reference and vice versa. In general, it is observed that the angular spectrum method ($d = 150$ mm) gives lower values for the cross-correlation, while the Fresnel diffraction method ($d = 500$ mm) gives the highest values. Also, the angular spectrum method requires the least amount of computations to generate the CGHs.

Figure 4-6(a) has been regenerated using less (28) and more (36) feature points. For 28 feature points, points 2,4,7 and 9 are eliminated; for 36, additional points are considered around the eyes. For both cases, the results are very similar to that in Figure 4-6(a), as shown in Figure 4-6(b) for the case of 28 feature points.

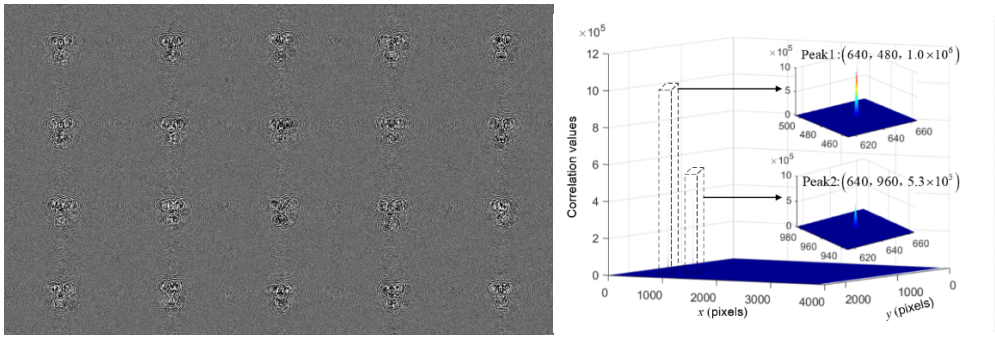


(a) (b)

Figure 4-6: The correlation peak values for verification of human faces. The dashed magenta line is a threshold for distinguishing the same face as the reference or not, which is equal to 4×10^5. (a) 32, (b) 28 feature points. [81]

As for facial identification, a reference face needs to be correlated to several faces simultanously, which is a typical 1-to-N problem. Here Face 1 is selected as the reference face and its CGH is the reference CGH. All the 20 CGHs are combined into a 4×5 cells, as shown in Figure 4-7(a), and correlated with the reference CGH. The result is shown in Figure 4-7(b). Two peak values are obvious as shown in Figure 4-7(b), Peak 1 denotes the autocorrelation peak value

for the reference CGH; Peak 2 is the correlation between Face 1 CGH and Face 6 CGH. Even though there are 18 correlation peak values in Figure 4-7(b), these values are too small to be visually identified. The identification using Face 1 as the reference is the same as the verification, previously shown in Figure 4-6(a).



(a) (b)

Figure 4-7: (a) CGHs for 20 faces by angular spectrum algorithm, mapping in 4×5 cells; (b) correlation coefficients for correlation between Figure 4-5(a) and Figure 4-7(a). Autocorrelation peak is labeled as Peak 1 and correlation peak between Face 1 and Face 6 is labeled as Peak 2. [81]

In order to verify the robustness of the proposed method, each of the 20 faces are taken as references, one at a time, and their CGHs are used to correlate with all CGHs. The results are shown in Figure 4-8. In Figure 4-8(a), the highest values are the autocorrelation results, the next highest values are the cross-correlations between the CGHs of Face 1 and Face 6, which are derived from the same person. The different colors and symbols denote the three different CGH algorithms with respective d's, as before. Once again, the dashed cyan line is a threshold for identification. This separation in peak values is sufficient for face recognition. It is apparent from Figure 4-8 that the cross-correlation peak values are mostly distributed in three regions. The black values obtained from Fresnel diffraction CGH algorithm ($d = 1000$ mm) are distributed between $8 \times 10^4 \sim 2 \times 10^5$; the red values for Rayleigh-Sommerfeld method ($d = 500$ mm) mostly fall in the range of $5 \times 10^4 \sim 1 \times 10^5$; and peak values of angular spectrum method ($d = 150$ mm) with blue markers are scattered approximately from $2 \times 10^4 \sim 5 \times 10^4$. The clusters of the correlation peak values reveal that CGH algorithms at different propagation distances perform distinctively.

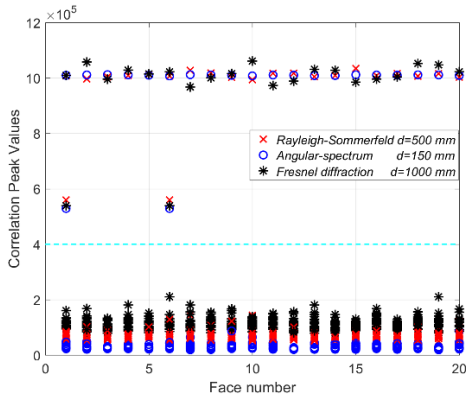
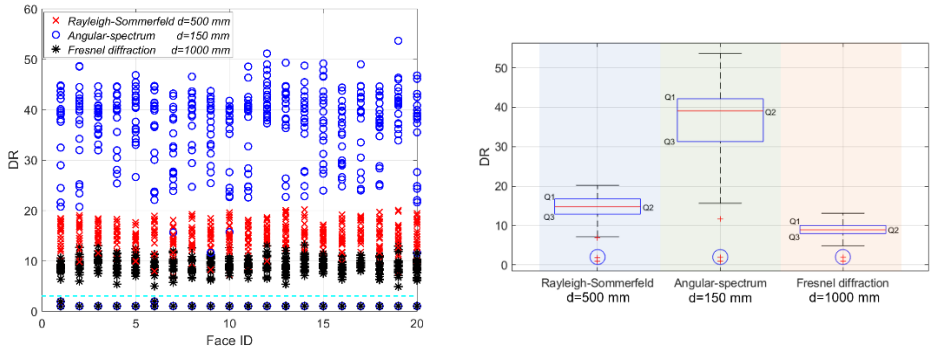


Figure 4-8: Correlation peak values for Face 1-20 sequentially being set as reference. A dashed cyan line is set as a threshold. [81]

4.2.2 Correlation Performance Evaluation

To further evaluate the correlation performance more quantitatively, the metrics introduced in Section 4.1 are calculated. The DR of Figure 4-8 is calculated, and results are shown in Figure 4-9(a). The higher value of DR denotes the larger discrimination between the reference and the target; small DR values or values close to one means that the reference and the target shares more similarities. In Figure 4-9(a), the three groups of colored cross stand for peak values obtained from three CGH algorithms at different propagation distances d, respectively. The DR for autocorrelation is equal to 1. A dashed cyan line with a value of 3 is drawn to separate DR values into two groups: for the values below the dashed line, the reference and the target are from the same person; otherwise, the target does not match the reference. A boxplot is drawn to analyze the correlation peak values in Figure 4-9(b). In Figure 4-9(b), for each CGH algorithm at their corresponding propagation distance, two horizontal solid lines are the maximum and minimum DR for cluster of values; the two solid blue horizontal lines and the solid red line are the three quartiles. For simplicity, the solid blue horizontal line with a smaller value is first quartile, labeled as Q1; the solid red line is the second quartile or median, labeled as Q2; similarly, the solid blue horizontal line with a larger value is labeled as Q3. For the values smaller than $Q_1 - 1.5 \times (Q_2 - Q_1)$ or larger than $Q_3 + 1.5 \times (Q_3 - Q_2)$, they are considered as outliers, which is shown in the red cross.

The blue circle showing the outliers corresponding to the values beneath the cyan dashed line in Figure 4-8(a). In Figure 4-9(b), the DR of angular spectrum is significantly higher than the other two methods. Upon comparing Rayleigh-Sommerfeld ($d = 500$ mm) and angular spectrum ($d = 150$ mm) methods, it is observed that they share a similar minimum value, but for DR values between Q1 and Q3, the Rayleigh-Sommerfeld shows a higher DR.

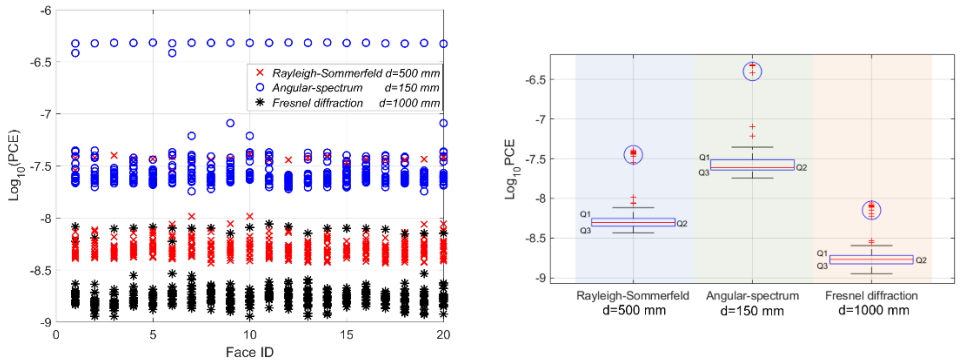


(a) (b)

Figure 4-9: (a) DR results for Figure 4-8(a). A dashed cyan line with a value of 3 is drawn to separate PCE values into two groups to evaluate whether the target face and reference face are obtained from the same person. Three colors and symbols denote three CGH algorithms. (b) Boxplot for (a). Black solid lines are the maximum and minimum values; blue horizontal lines are the first and third quartiles labeled as Q1 and Q3; red solid line is the median, labeled as Q2; the red cross are the values smaller than $Q_1 - 1.5 \times (Q_2 - Q_1)$ or larger than $Q_3 + 1.5 \times (Q_3 - Q_2)$; blue circle includes the values beneath the dashed cyan line in (a). [81]

The PCE measures how much energy is accumulated in the correlation peak, or in other words, represents the sharpness of the correlation peak. For easier comparison, a logarithm is performed on PCE values. An obvious separation is shown in Figure 4-10(a): correlation of CGHs generated by angular spectrum with $d = 150$ mm shows larger PCE values than the other two methods. Figure 4-10(b) gives the statistical result for Figure 4-10(a). For angular spectrum and Fresnel diffraction, correlation peak values of angular spectrum with $d = 150$ mm are roughly 4 times larger than the Fresnel diffraction with $d = 1000$ mm, however, the PCE of angular spectrum method is more than 10 times larger than the Fresnel diffraction method. The results indicate that the angular spectrum algorithm shows a better correlation peak sharpness compared to Fresnel

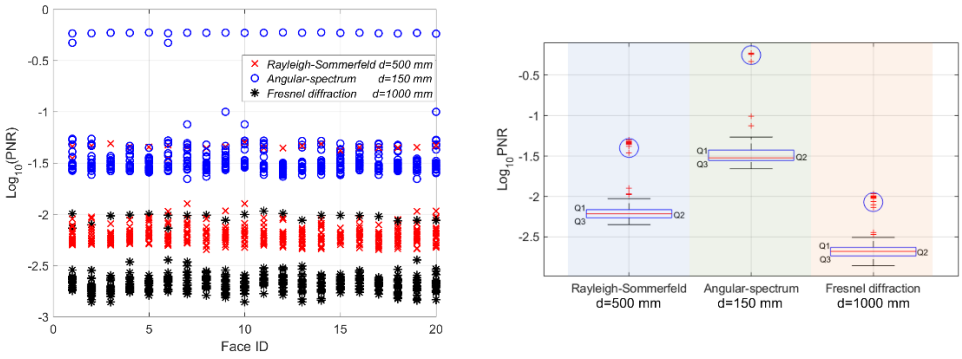
diffraction. A similar comparison can be done between the angular spectrum method and Rayleigh-Sommerfeld method.



(a) (b)

Figure 4-10: (a) $\log_{10}$PCE results for Figure 4-8(a). (b) Boxplot for (a), showing a statistic analysis of $\log_{10}$PCE. [81]

Finally, the PNR indicates the correlation peak value divided by the mean of correlation plane energy. It is always preferred that the correlation peak is much higher than the mean of surrounding noise. Results presented in Figure 4-11 give a straightforward comparison of PNR for three methods, and show that angular spectrum method shows a higher PNR than the other two methods, which are consistent to the DR and PCE results.



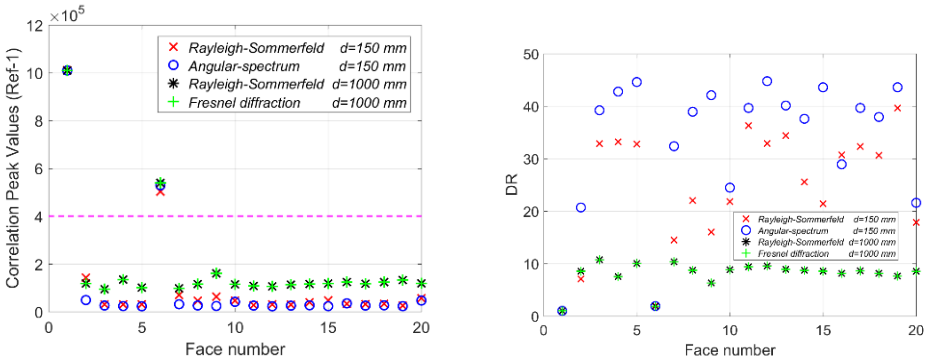
(a) (b)

Figure 4-11: (a) $\log_{10}$PNR results for Figure 4-8(a). (b) Boxplot for (a), showing a statistic analysis of $\log_{10}$PNR. [81]

Investigation of figures of merits shows that angular spectrum layer-oriented method with $d = 150$ mm is superior to the other two methods in CGH correlation performance. By using the angular spectrum method, the correlation peak values give more separation between the reference human face and the target human face, and the correlation plane energy is much smaller with more energy accumulated in correlation peak values. Among these three methods, the angular spectrum method ($d = 150$ mm) encoding of 3D face feature information into a 2D CGH, can be considered as the best candidate for generation of CGHs to be correlated for face recognition.

Finally, the effect of propagation distance on the results obtained using the three methods is investigated. Since Rayleigh-Sommerfeld method has no strict restriction on propagation distance, CGHs are generated by this method at $d = 150$ mm and 1000 mm. Correlation results obtained using CGHs generated at $d = 150$ mm and $d = 1000$ mm are compared with correlation results from the Fresnel method (at $d = 1000$ mm) and the angular spectrum method (at $d = 150$ mm). The results are shown in Figure 4-12(a). It is clear from the figure that overall the Rayleigh-Sommerfeld method for $d = 1000$ mm agrees well with the Fresnel method for the same distance; and the Rayleigh-Sommerfeld method for $d = 150$ mm has reasonably good agreement with the angular spectrum method for the same distance. A longer propagation distance may lead to the exclusion of some distinguishing higher spatial frequency information in the holograms, which in turn can give rise to higher cross-correlation peaks and may lead to a lower discrimination between CGH templates. The DRs obtained using the angular spectrum method ($d = 150$ mm), Rayleigh-Sommerfeld method (for $d = 150$ mm), Fresnel method ($d = 1000$ mm), and the Rayleigh-Sommerfeld method (for $d = 1000$ mm) are shown in Figure 4-12(b). While the DRs for the Rayleigh-Sommerfeld method and the Fresnel method for $d = 1000$ mm are very similar, there are noticeable differences between the DRs using the Rayleigh-Sommerfeld method and the angular spectrum method for $d = 150$ mm. These differences possibly come from the slightly higher values for the cross-correlation observed in the case of Rayleigh-Sommerfeld compared to

the angular spectrum method. This may be due to the fact that in the Rayleigh-Sommerfeld formulation taken for computations of the CGHs (see Eq. (4-2)), evanescent solutions are not included (and neither is the obliquity factor), whereas the angular spectrum approach derived directly from the Helmholtz equation using Fourier transformation) has no such assumptions. The conversion from the transfer function to the impulse response which is used for the Rayleigh-Sommerfeld method and the assumptions involved are discussed in detail in Stark [102] and Poon and Banerjee [103]. The DRs for $d = 150$ mm are higher than those at $d = 1000$ mm, for reasons discussed above.



(a) (b)

Figure 4-12: (a) Correlation peak values for recognition of human faces with $d = 150$ mm for Rayleigh-Sommerfeld and angular spectrum, $d = 1000$ mm for Rayleigh-Sommerfeld and Fresnel diffraction; (b) DR results for (a). [81]

4.3 Conclusion and Discussion

We have proposed a CGH correlation method for face recognition, providing a new way for 3D face recognition. Three-dimensional human faces are pre-processed and feature points are extracted for further evaluation. The CGHs corresponding to the feature points are generated by three algorithms: Rayleigh-Sommerfeld point-oriented algorithm, angular spectrum algorithm and Fresnel approximation, respectively. Correlations are performed on reference and target CGHs, and peak values are extracted for face verification and identification. Faces of the same person with different facial expressions have been recognized and categorized; while, faces of different

individuals have been detected and identified. Correlation performance metrics, viz. DR, PCE, and PNR, are computed to evaluate the quality of correlation results with different CGH algorithms. These figures of merits indicate the angular spectrum layer-oriented method with $d = 150$ mm performs better than the other two methods for face recognition. A shorter propagation distance for preservation of higher spatial frequency information in the angular spectrum-based layer-oriented technique shows the best performance. Furthermore, this method is computationally the fastest. For high-throughput application of this method in 3D face recognition, further work still needs to be done. In the future, we will investigate the optimal parameters of feature points and suitable range of all the parameters for facial recognition. Automated preprocessing of the 3D models and extraction of the sampling points is necessary for a large database and could be useful in commercial/industrial applications; this will also be examined in future research. Currently, the robustness of our method is dependent on selection of feature points. However, the proposed method holds good promise for 3D facial recognition, with the advantage in encoding and decoding, possible combination with existing optical and numerical image recognition techniques, and hopefully reduction in computational effort.

CHAPTER V

CONCLUSIONS AND FUTURE WORK

5.1 Conclusions

In this thesis, phase retrieval using TIE with TPE, phase reconstruction using DH with TIE, and correlation of holograms are introduced.

In Chapter II, a non-iterative FFT-based TIE+TPE phase retrieval method is proposed. The method takes the diffraction effect into account by incorporating non-paraxial TPE. The technique can achieve better accuracy than traditional FFT-based TIE method with little sacrifice in computational time, which has been shown numerically and experimentally.

In Chapter III, the DH+TIE method is described. The performance of DH+TIE, SWDH, and DWDH are evaluated using MSE and computational time. Specifically, SWDH and DWDH using PUMA for phase unwrapping have been compared with DH+TIE in both in-line and off-axis geometry, along with a newly introduced PSDH+TIE technique. CGHs as well as experimentally recorded digital holograms have been conducted in phase retrieval. The DH+TIE method shows advantage in automatic multiple image registration in numerical reconstruction process; single capture of the hologram in achieving phase retrieval without any phase unwrapping process; no moving components in the optical system are needed; and much shorter computational time compared to traditional DH methods.

In Chapter IV, a CGH correlation method for face recognition provides a new way for 3D face recognition. CGHs are generated from feature points of the human face data through optical propagation theories. Peak values of correlations between reference and target CGHs are extracted for face verification and identification. From the peak values, faces of the same person with different facial expressions are recognized and categorized; while, faces obtained different

individuals are verified and identified. Figures of merits are calculated to show the quality of the correlation results, which indicate that the usage of the angular spectrum-based layer-oriented technique gives the best correlation quality.

Chapter II's work is presented at *SPIE Photonics West 2021 Digital Forum* and accepted by *Proceedings of SPIE* [104]; Chapter III's work is published in *Applied Optics* [79]; Chapter IV's work is published in *Applied Optics* [81].

5.2 Future Work

In the future, TIE+TPE method can be potentially applied to real-time phase imaging and amplitude-phase imaging. Improvement can be made to achieve single-shot imaging by using grating in the system or combing with digital holography. DH+TIE method can be developed with a variety of TIE solvers in achieving better accuracy. Also, partially coherent light can be introduced to DH+TIE method to reduce or avoid speckle noises. For correlation of computer-generated holograms in face recognition, further optimization can be made by introducing deep learning to the propagation modeling to increase the robustness of the system.

REFERENCES

1. G. Popescu, *Quantitative Phase Imaging of Cells and Tissues*. McGraw-Hill (2011).
2. C. L. Curl, C. J. Bellair, P. J. Harris, B. E. Allman, A. Roberts, K. A. Nugent, and L. M. D. Delbridge, "Quantitative phase microscopy: a new tool for investigating the structure and function of unstained live cells," Clin. Exp. Pharmacol. Physiol. **31**, 896–901 (2004).
3. C. L. Curl, C. J. Bellair, P. J. Harris, B. E. Allman, A. Roberts, K. A. Nugent, L. M. D. Delbridge, "Single cell volume measurement by quantitative phase microscopy (QPM): a case study of erythrocyte morphology," Cell Physiol. Biochem. **17**, 193–200 (2006).
4. H. Guo, H. Zhou, and P. P. Banerjee, "Single -shot digital phase-shifting Moiré patterns for 3D topography," Appl. Opt. **60**, A84-A92 (2020).
5. B. Bordbar, M. M. R. Hussain, and P. P. Banerjee, "Application of complex field imaging sensor to additive manufacturing," Proc. of SPIE, **11305**, 113050G (2020).
6. U. Abeywickrema, P. P. Banerjee, A. Kota, S. Swiontek, and A. Lakhtakia, "High-resolution topograms of fingerprints using multiwavelength digital holography," Opt. Eng. **56**, 034117-1-8 (2017).
7. F. Tsalakanidou, F. Forster, S. Malassiotis, M. G. Strintzis, "Real-time acquisition of depth and color images using structured light and its application to 3D face recognition," Real-Time Imaging **11**, 358-369 (2005).
8. F. Zernike, "Phase contrast, a new method for the microscopic observation of transparent objects," Physica **9**, 974–986 (1942).
9. G. M. Nomarski, "Differential microinterferometer with polarized waves," J. Phys. Radium, **16**, 9S-11S (1955).
10. B. P. Abbott *et al*., "GW170817: Observation of gravitational waves from a binary neutron star inspiral," Phys. Rev. Lett. **119**, 161101-1-18 (2017).

11. D. Gabor, “A new microscopic principle,” Nature **161**, 777-778, (1948).
12. J. Goodman and R. Lawrence, “Digital image formation from electronically detected holograms,” Appl. Phys. Lett. **11**, 77-79, (1967).
13. J. Fienup, “Phase retrieval algorithms: a comparison,” Appl. Optics **21**, 2758-2769 (1982).
14. R. W. Gerchberg and W. O. Saxton, “A practical algorithm for the determination of the phase from image and diffraction plane pictures,” Optik **35**, 237-246 (1972).
15. J. Miao, P. Charalambous, J. Kirz, and D. Sayre, “Extending the methodology of X-ray crystallography to allow imaging of micrometer-sized non-crystalline specimens,” Nature **400**, 342-344 (1999).
16. B. Bordbar, M. M. R. Hussain, and Partha P. Banerjee, “Application of complex field imaging sensor to additive manufacturing,” Ultra-High-Definition Imaging Systems III **11305**, 11305G (2020).
17. R. Gonsalves and R. Chidlaw, “Wavefront sensing by phase retrieval,” Applications of Digital Image Processing III **207**, 32-39 (1979).
18. O. Guyon, “Limits of adaptive optics for high-contrast imaging,” Astrophys. J. **629**, 592 (2005).
19. H. Wang, W. Yue, Q. Song, J. Liu, and G. Situ, “A hybrid Gerchberg-Saxton-like algorithm for DOE and CGH,” Opt. Lasers Eng. **89**, 109-115 (2017).
20. H. M. L. Faulkner and J. M. Rodenburg, “Moveable aperture lensless transmission microscopy: a novel phase retrieval algorithm,” Phys. Rev. Lett. **93**, 023903-1-4 (2004).
21. H. M. L. Faulkner and J. M. Rodenburg, “Error tolerance of an iterative phase retrieval algorithm for moveable illumination microscopy,” Ultramicroscopy **103**, 153-164 (2005).
22. G. Zheng, R. Horstmeyer, and C. Yang, “Wide-field, high-resolution Fourier ptychographic microscopy,” Nat. Photonics **7**, 739-745 (2013).

23. C. Zuo, J. Li, J. Sun, Y. Fan, J. Zhang, L. Lu, R. Zhang, B. Wang, L. Huang, and Q. Chen, "Transport of intensity equation: a tutorial," Opt. Lasers Eng. **135**, 106187 (2020).
24. M. R. Teague, "Irradiance moments: their propagation and use for unique retrieval of phase," J. Opt. Soc. Am. **72**, 1199–209 (1982).
25. M. R. Teague, "Deterministic phase retrieval: a Green's function solution," J. Opt. Soc. Am. A **73**, 1434-1441 (1983).
26. P. P. Banerjee, A. Korpel, and K. E. Lonngren, "Self-refraction of nonlinear capillary-gravity waves," Phys. Fluids **26**, 2393-2398 (1983).
27. P. P. Banerjee and T.-C. Poon, *Principles of Applied Optics.* CRC Press (1991).
28. S. A. Akhmanov, A. P. Sokhorukov, and R. V. Khokhlov, "Self-focusing and self-trapping of intense light beams in a nonlinear medium," Sov. Phys. JETP **23**, 1025-1033 (1966).
29. T. E. Gureyev, A. Roberts, and K. A. Nugent, "Partially coherent fields, the transport-of-intensity equation, and phase uniqueness," J. Opt. Soc. Am. A **12**, 1942-1946 (1995).
30. T. E. Gureyev, "Composite techniques for phase retrieval in the Fresnel region," Opt. Commun. **220**, 49-58 (2003).
31. T. E. Gureyev, A. Roberts, and K. A. Nugent, "Phase retrieval with the transport-of-intensity equation: matrix solution with use of Zernike polynomials," J. Opt. Soc. Am. A **12**, 1932-1941 (1995).
32. G. Yang, B. Dong, B. Gu, J. Zhuang, and O. Ersoy, "Gerchberg–Saxton and Yang–Gu algorithms for phase retrieval in a nonunitary transform system: a comparison," Appl. Opt. **33**, 209-218 (1994).
33. F. Roddier and C. Roddier. "Wavefront reconstruction using iterative Fourier transforms," Appl. Opt. **30**, 1325–1327 (1991).
34. L. J. Allen and M. P. Oxley, "Phase retrieval from series of images obtained by defocus variation," Opt. Commun. **199**, 65-75 (2001).

35. C. Zuo, Q. Chen, and A. Asundi, “Boundary-artifact-free phase retrieval with the transport of intensity equation: fast solution with use of discrete cosine transform,” Opt. Exp. **22**, 9220-9244 (2014).

36. M. Basunia, P. P. Banerjee, U. Abeywickrema, T.-C. Poon, and H. Zhang, “Recursive method for phase retrieval using transport of intensity and its applications,” Appl. Opt. **55**, 9546–9554 (2016).

37. H. Zhang, W. Zhou, Y. Liu, T. Tian, P. Banerjee, and T-C. Poon, “Finite difference approach to transport of intensity,” Digital Holography and Three-Dimensional Imaging, OSA Technical Digest (online) W1A.5 (2017).

38. H. Zhang, W-J. Zhou, Y. Liu, D. Leber, P. Banerjee, M. Basunia, and T.-C. Poon, “Evaluation of finite difference and FFT-based solutions of the transport of intensity equation,” Appl. Opt. **57**, A222-A228 (2018).

39. D. Paganin, and K. Nugent, “Non-interferometric phase imaging with partially coherent light,” Phys. Rev. Lett. **80**, 2586–2589 (1998).

40. K. Nugent, D. Paganin, and T. Gureyev, "A phase odyssey," Physics Today, **54**, 27-32 (2001).

41. B. E. Allman, P. J. McMahon, K. Nugent, D. Paganin, D. L. Jacobson, M. Arif, and S. A. Werner, “Phase radiography with neutrons,” Nature **408**, 158-159 (2000).

42. P. J. McMahon, B. E. Allman, D. L. Jacobson, M. Arif, S. A. Werner, K. Nugent, “Quantitative phase radiography with polychromatic neutrons,” Phys. Rev. Lett. **91**, 145502 (2003).

43. K. Nugent, T. E. Gureyev, D. Cookson, D. Paganin, and Z. Barnea, “Quantitative phase imaging using hard x rays,” Phy. Rev. Lett. **77**, 2961-2964 (1996).

44. C. Zuo, Q. Chen, W. Qu, and A. Asundi, “High-speed transport-of-intensity phase microscopy with an electrically tunable lens,” Opt. Exp. **21**, 24060-24075 (2013).

45. L. Waller, Y. Luo, S. Y. Yang, and G. Barbatathis, “Transport of intensity phase imaging in a volume holographic microscope,” Opt. Lett. **35**, 2961-2963 (2010).
46. U. Schnars and W. Juptner, *Digital Holography.* Springer (2005).
47. G. Nehmetallah and P. P. Banerjee, “Applications of digital and analog holography in three-dimensional imaging,” Adv. Opt. Photon. **4**, 472–553 (2012).
48. T.-C. Poon and J.-P. Liu, *Introduction to Modern Digital Holography with MATLAB.* Cambridge University Press (2014).
49. G. Nehmetallah, R. Aylo, and L. Williams, *Analog and Digital Holography with MATLAB,* SPIE (2015).
50. I. Yamaguchi and T. Zhang, "Phase-shifting digital holography," Opt. Lett. **22**, 1268-1270 (1997).
51. J. M. Bioucas-Dias and G. Valadao, “Phase unwrapping via graph cuts,” IEEE Trans. Image Process **16**, 698-709 (2007).
52. N. Pandey, A. Ghosh, and K. Khare, “Two-dimensional phase unwrapping using the transport of intensity equation,” Appl. Opt. **55**, 2418-2425 (2016).
53. J. Martinez, K. Falaggis, and T. Kozacki, “Fast and accurate phase-unwrapping algorithm based on the transport of intensity equation,” Appl. Opt. **56**, 7079-7088 (2017).
54. Z. Zhao, H. Zhang, Z. Xiao, H. Du, Y. Zhuang, C. Fan, and H. Zhao, “Robust 2D phase unwrapping algorithm based on the transport of intensity equation,” Meas. Sci. Technol. **30**, 015201 1-8 (2019).
55. R. Horisaki, R. Egami, and J. Tanida, “Single-shot phase imaging with randomized light (SPIRaL),” Opt. Exp. **24**, 3765-3773 (2016).
56. R. Horisaki, Y. Ogura, M. Aino, and J. Tanida, “Single-shot phase imaging with a coded aperture,” Opt. Lett. **39**, 6466-6469 (2014).

57. B. Bordbar, M. M. R. Hussain, and P. P. Banerjee, “Application of complex field imaging sensor to additive manufacturing,” Proc. of SPIE **11305**, 113050G (2020).

58. C. Zuo, Q. Chen, W. Qu, and A. Asundi, “Direct continuous phase demodulation in digital holography with use of the transport-of-intensity equation,” Opt. Commun. **309**, 221-226 (2013).

59. W. Zhou, X. Guan, F. Liu, Y. Yu, H. Zhang, TC. Poon, and P.P. Banerjee, “Phase retrieval based on transport of intensity and digital holography,” Appl. Opt. **57**, A229-A234 (2018).

60. W. Zhou, H. Shen, X. Guan, Y. Yu, H. Zhang, and TC-Poon, “Simulation analysis on phase retrieval using transport of intensity with an off-axis hologram”, IEEE 28th International Symposium on Industrial Electronics (ISIE), 2419-2424, (2019).

61. A. F. Abate, M. Nappi, D. Riccio, and G. Sabatino, “2D and 3D face recognition: A survey”, Pattern Recognition Lett. **28**, 1885-1906 (2007).

62. K. W. Bowyer, K. Chang, and P. Flynn, “A survey of approaches and challenges in 3D and multi-modal 3D + 2D face recognition”, Computer Vision and Image Understanding **101**, 1-15 (2006).

63. C. Xu, Y. Wang, T. Tan, and L. Quan, “Depth vs. intensity: which is more important for face recognition?” Proc. of 17th International Conf. on Pattern Recognition **4**, 342–345 (2004).

64. Y. Liu, X. Cao, Q. Dai and W. Xu, “Continuous depth estimation for multi-view stereo”, IEEE International Conf. Computer Vision and Pattern Recognition 2121-2128 (2009).

65. V. Blanz and T. Vetter, “Face recognition based on fitting a 3D morphable model”, IEEE Transactions on Pattern Analysis and Machine Intelligence **25**, 1063-1074 (2003).

66. R. O. Duda, P. E. Hart, and D. G. Stork*, Pattern Classification*. Wiley (2001).

67. A. Ansari and M. Abdel-Mottaleb, “3-D face modeling using two views and a generic face model with application to 3-D face recognition”, Proc. of the IEEE Conf. on Advanced Video and Signal Based Surveillance 37-44 (2003).

68. K. Chang, K. Bowyer, and P. Flynn, “Face recognition using 2D and 3D facial data”, Proc. of ACM Workshop on Multimodal User Authentication 25–32 (2003).

69. K. Chang, K. W. Bowyer, and P. Flynn, “Multi-modal 2D and 3D biometrics for face recognition”, Proc. of the IEEE International Workshop on Analysis and Modeling of Faces and Gestures 187-194 (2003).

70. X. Tu, J. Zhao, Z. Jiang, Y. Luo, M. Xie, Y. Zhao, L. He, Z. Ma, and J. Feng, “Joint 3D face reconstruction and dense face alignment from a single image with 2D-assisted self-supervised learning”, IEEE International Conf. Computer Vision and Pattern Recognition (2019).

71. A. Jourabloo and X. Liu, “Large-pose face alignment via CNN-based dense 3D model fitting”, IEEE International Conf. Computer Vision and Pattern Recognition (2016).

72. E. Watanabe and K. Kodate, "Implementation of a high-speed face recognition system that uses an optical parallel correlator," Appl. Opt. **44**, 666-676 (2005).

73. U. Abeywickrema, R. Gnawali, and P.P. Banerjee, “Identification of 3D objects using correlation of holograms,” Proc. SPIE **10752**, 1075219 (2018).

74. H. Zhou, U. Abeywickrema, B. Bordbar, L. Cao, and P.P. Banerjee, “Correlation of holograms for surface characterization of diffuse objects,” Proc. SPIE **10943**, 10943-3 (2019).

75. H. Zhou, R. Hou, B. Bordbar, and P.P. Banerjee, “Effect of hologram windowing on correlation of 3D objects,” Digital Holography and 3-D Imaging OSA Th2B.8 (2019).

76. T. Kim and T.-C. Poon, “Three-dimensional matching by use of phase only holographic information and the Wigner distribution,” J. Opt. Soc. Amer. A **17**, 2520-2528 (2000).

77. Y. Yi and L. Cao, “Optical fingerprint recognition based on local minutiae structure coding”, Opt. Express **21**, 17108-17121 (2013).

78. H. Zhang, Y. Zhao, L. Cao, and G. Jin, “Fully computed holographic stereogram based algorithm for computer-generated holograms with accurate depth cues”, Opt. Express **23**, 3901-3913 (2015).

79. H. Zhou, E. Stoykova, M. Hussain, and P. P. Banerjee, “Performance analysis of phase retrieval using transport of intensity with digital holography,” *invited paper*, Appl. Opt. **60**, A73-A83 (2020).

80. L. Waller, L. Tian, and G. Barbastathis, “Transport of intensity phase-amplitude imaging with higher order intensity derivatives,” Opt. Exp. **18**, 12553-12561 (2010).

81. H. Zhou, X. Sui, L. Cao, and P. P. Banerjee, “Digital correlation of computer-generated holograms for 3D face recognition,” Appl. Opt. **58**, G177-G186 (2019).

82. S. Agarwal, V. Kumar, and C. Shakher, “Onion epidermis cell imaging by using Talbot effect”, Imaging and Applied Optics JTu5A.27 (2017).

83. B. Bordbar, H. Zhou, and P. P. Banerjee, “3D object recognition through processing of 2D holograms,” Appl. Opt. **58**, G177-G186 (2019).

84. T.-C. Poon, *Optical Scanning Holography with MATLAB*, Springer (2007).

85. G. Brooker, N. Siegel, J. Rosen, N. Hashimoto, M. Kurihara, and A. Tanabe, “In-line FINCH super resolution digital holographic fluorescence microscopy using a high efficiency transmission liquid crystal GRIN lens,” Opt. Lett. **38**, 5264–5267 (2013).

86. Y. Ichihashi, H. Nakayama, T. Ito, N. Masuda, T. Shimobaba, A. Shiraki, and T. Sugie, “HORN-6 special-purpose clustered computing system for electroholography,” Opt. Express **17**, 13895–13903 (2009).

87. A. Shiraki, N. Takada, M. Niwa, Y. Ichihashi, T. Shimobaba, N. Masuda, and T. Ito, "Simplified electroholographic color reconstruction system using graphics processing unit and liquid crystal display projector," Opt. Express **17**, 16038–16045 (2009).

88. Y. Zhao, L. Cao, H. Zhang, and Q. He, "Holographic display with LED illumination based on phase only spatial light modulator," Proc. SPIE **8559**, 85590B (2012).

89. Y. Zhao, L. Cao, H. Zhang, D. Kong, and G. Jin, "Accurate calculation of computer-generated holograms using angular-spectrum layer-oriented method", Opt. Express **23**, 25440-25449 (2015).

90. O. Wolfgang and N. Reingand, *Optical Imaging and Metrology: Advanced Technologies*. Wiley (2012).

91. H. Zhang, L. Cao, and G. Jin, "Computer-generated hologram with occlusion effect using layer-based processing", Appl. Opt. **56**, F138-F143 (2017).

92. Y. Zhao, L. Cao, H. Zhang, W. Tan, S. Wu, Z. Wang, Q. Yang, and G. Jin, "Time-division multiplexing holographic display using angular-spectrum layer-oriented method", Chinese Opt. Lett. **14**, 010005-1-5 (2016).

93. P. P. Banerjee, U. Abeywickrema, H. Zhou, B. Bordbar, M. Alam, G. Nehmetallah, J. Khoury, and L. Cao, "Taking correlation from 2D to 3D: Optical methods and performance evaluation," Proc. SPIE **10995**, (2019).

94. G. Nehmetallah, J. Khoury, M. A. Alam, and P.P. Banerjee, "Photorefractive two-beam coupling joint transform correlator: modeling and performance evaluation," Appl. Opt. **55**, 4011-4023 (2016).

95. A. Awwal, K. Gudmundsson, M. Tabrez, M. Rahman, M. Alam, and K. Iftekharuddin, "A new metric for 3D optical pattern recognition," Proc. SPIE **4788**, 183–190 (2002).

96. M. Alam, A. Bal, E. Horache, S. Goh, C. Loo, S. Regula, and A. Sharma, "Metrics for evaluating the performance of joint transform correlation based target recognition and tracking algorithms," Opt. Eng. **44**, 067005 (2005).

97. J. L. Horner, "Metrics for assessing pattern-recognition performance," Appl. Opt. **31**, 165–166 (1992).

98. G. Nehmetallah, P. P. Banerjee, M. Alam, and J. Khoury, "Performance evaluation of photorefractive two-beam coupling joint transform correlator," Proc. SPIE **9094**, 909409 (2014).

99. J. Khoury, M. S. Alam, P. P. Banerjee, G. Nehmetallah, W. M. Durant, D. M. Martin, J. Donoghue, N. Peyghambarian, and M. Yamamoto, "Performance comparison of photorefractive two-beam coupling correlator with optimal filter based correlators," Proc. SPIE **9094**, 909405 (2014).

100. G. S. Pati, R. Tripathi, and K. Singh, "Photorefractive joint-transform correlator using incoherent- erasure in two-wave-mixing geometry," Opt. Commun. **151**, 268–272 (1998).

101. J. D. Schmidt. *Numerical Simulation of Optical Wave Propagation with Examples in MATLAB*, Bellingham, Washington USA, SPIE (2010).

102. H. Stark ed., *Applications of Optical Transforms*, Academic, New York (1982).

103. T.-C. Poon and P. P. Banerjee, *Contemporary Optical Image Processing with MATLAB*, Elsevier, Amsterdam (2001); Chapter 3 pp. 54-56.

104. H. Zhou and P. P. Banerjee, "Transport of intensity phase imaging with error correction using transport of phase equation," Proc. SPIE **11709**, 117090D (2021).

APPENDIX A

Derivation of TIE and TPE

The starting point of the derivation of TIE is the Helmholtz equation for the scalar optical field

$$\nabla^2 E_p + {k_0}^2 E_p = 0, \tag{A-1}$$

where k_0 is the propagation constant, and E_p stands for the optical field phasor,

$$E_p(\vec{r}_\perp, z) = E_e(\vec{r}_\perp, z)e^{-jk_0 z}, \tag{A-2}$$

with E_e representing the envelope of the optical field. $\vec{r}_\perp$ represents the transverse coordinates (x, y), and z-axis the light propagation direction. We can represent the optical field phasor and envelope as

$$E_p(\vec{r}_\perp, z) = a(\vec{r}_\perp, z)e^{-j\tilde{\phi}(\vec{r}_\perp, z)}, E_p(\vec{r}_\perp, z) = a(\vec{r}_\perp, z)e^{-j\phi(\vec{r}_\perp, z)}, \tag{A-3}$$

where $a(\vec{r}_\perp, z) = \sqrt{I(\vec{r}_\perp, z)}$ denotes the amplitude with I denoting the intensity. It is noted that $\tilde{\phi}(\vec{r}_\perp, z)$ represents the phase of E_p and is related to the phase $\phi(\vec{r}_\perp, z)$ of E_e through the relation $\tilde{\phi} = \phi + k_0 z$.

Substituting Eq. (A-3) into Eq. (A-1),

$$\nabla^2\left(ae^{-j\tilde{\phi}}\right) + {k_0}^2\left(ae^{-j\tilde{\phi}}\right) = 0. \tag{A-4}$$

Then, using the operator identities: (a) $\boldsymbol{\nabla} \cdot \left(a\vec{A}\right) = \boldsymbol{\nabla} a \cdot \vec{A} + a\boldsymbol{\nabla} \cdot \vec{A}$, (b) $\boldsymbol{\nabla}(ab) = b\boldsymbol{\nabla} a + a\boldsymbol{\nabla} b$, Eq. (A-4) can be re-expressed as:

$$\left(e^{-j\tilde{\phi}}\right)\nabla^2 a + 2\boldsymbol{\nabla}\left(e^{-j\tilde{\phi}}\right) \cdot \boldsymbol{\nabla} a + a\nabla^2\left(e^{-j\tilde{\phi}}\right) + {k_0}^2\left(ae^{-j\tilde{\phi}}\right) = 0. \tag{A-5}$$

Equation (A-5) can be further simplified to show the relation directly between the phase $\tilde{\varphi}$ and the amplitude a:

$$\left(e^{-j\tilde{\phi}}\right)\nabla^2 a - 2j\left(e^{-j\tilde{\phi}}\right)\boldsymbol{\nabla}\tilde{\phi}\cdot\boldsymbol{\nabla}a - \left(e^{-j\tilde{\phi}}\right)a\boldsymbol{\nabla}\tilde{\phi}\cdot\boldsymbol{\nabla}\tilde{\phi} - j\left(e^{-j\tilde{\phi}}\right)a\nabla^2\tilde{\phi} + {k_0}^2 a\left(e^{-j\tilde{\phi}}\right) = 0. \tag{A-6}$$

From the imaginary and real part in Eq. (A-6), two equations can be obtained:

$$2j\boldsymbol{\nabla}\tilde{\phi}\cdot\boldsymbol{\nabla}a + ja\nabla^2\tilde{\phi} = 0, \tag{A-7}$$

$$\nabla^2 a - a\boldsymbol{\nabla}\tilde{\phi}\cdot\boldsymbol{\nabla}\tilde{\phi} + {k_0}^2 a = 0. \tag{A-8}$$

Multiplying Eq. (A-7) by a, using the relation $2a\boldsymbol{\nabla}a = \boldsymbol{\nabla}a^2$ and the vector identity $\boldsymbol{\nabla}\cdot\left(a\vec{A}\right) = \boldsymbol{\nabla}a\cdot\vec{A} + a\boldsymbol{\nabla}\cdot\vec{A}$ once again, Eq. (A-7) can be expressed as:

$$\boldsymbol{\nabla}\cdot\left(a^2\boldsymbol{\nabla}\tilde{\phi}\right) = 0. \tag{A-9}$$

This is the non-paraxial form of the TIE, derived from the imaginary part of the Helmholtz equation. For the real part, with some algebra, Eq. (A-8) can be expressed as

$$\boldsymbol{\nabla}\tilde{\phi}\cdot\boldsymbol{\nabla}\tilde{\phi} = \frac{1}{a}\nabla^2 a + {k_0}^2. \tag{A-10}$$

This is the non-paraxial form of the TPE, which is similar to the eikonal equation in the presence of the diffraction term $\frac{1}{a}\nabla^2 a$.

Even though Eqs. (A-9) and (A-10) are very compact, they are hard to solve or implement. In common scenarios, paraxial approximations are invoked. Using the relation $\tilde{\phi} = \phi + k_0 z$, Eqs. (A-9) and (A-10) become:

$$\boldsymbol{\nabla}_\perp\cdot\left(a^2\boldsymbol{\nabla}_\perp\phi\right) + \frac{\partial}{\partial z}\left(a^2\left(\frac{\partial\phi}{\partial z} + k_0\right)\right) = 0, \tag{A-11}$$

$$\boldsymbol{\nabla}_\perp\phi\cdot\boldsymbol{\nabla}_\perp\phi + 2k_0\left(\frac{\partial\phi}{\partial z}\right) + \left(\frac{\partial\phi}{\partial z}\right)^2 = \frac{1}{a}\left(\nabla_\perp^2 a + \frac{\partial^2 a}{\partial z^2}\right). \tag{A-12}$$

Under the condition that $\frac{\partial \phi}{\partial z} \ll k_0$, Eq. (A-11) can be approximated to

$$\boldsymbol{\nabla}_\perp \cdot (a^2 \boldsymbol{\nabla}_\perp \phi) + k_0 \frac{\partial a^2}{\partial z} = 0. \tag{A-13}$$

Since $a^2 = I$, Eq. (A-13) is the same as Eq. (2-6), the paraxial form of the TIE. Also, using the above condition and the assumption, $\frac{\partial^2 a}{\partial z^2} \ll \nabla_\perp^2 a$, we can also obtain the paraxial form of the TPE:

$$\boldsymbol{\nabla}_\perp \phi \cdot \boldsymbol{\nabla}_\perp \phi + 2k_0 \left(\frac{\partial \phi}{\partial z}\right) = \frac{1}{a} \nabla_\perp^2 a. \tag{A-14}$$

APPENDIX B

MATLAB code for DH+TIE simulation

```
%% off-axis DH+TIE Simulation Code
% Written by Haowen Zhou
% ============================================================
close all;clear all;clc;


%% parameters
Nx = 1024; Ny = 1024;     % number of pixels along X and Y axes
lambda1 = 514.5e-9;       % wavelength in m
k0 = 2*pi/lambda1;        % wave number
dz = 1e-4;                % two adjacent intensity seaparation along z axis
dzm = -dz:dz:dz;          % delta z in m
dx = 4.65e-6;             % pixel pitch in m
dy = dx;                  % pixel pitch in y direction is the same as x-axis
d = 20e-2;                % obj to ccd distance in m
x_obj = linspace(-dx*Nx/2,dx*Nx/2,Nx);        % row for ccd plane
y_obj = linspace(-dy*Ny/2,dy*Ny/2,Ny);        % col for ccd plane
[X, Y] = meshgrid(y_obj,x_obj);               % meshgrid
kx = -pi/dx: 2*pi/dx/Nx: pi/dx-2*pi/dx/Nx;    % kx axis
ky = -pi/dy: 2*pi/dy/Ny: pi/dy-2*pi/dy/Ny;    % ky axis
[Kx,Ky] = meshgrid(ky, kx);                   % kx and ky are transverse spatial
frequency


%% Generation of a phase object
Pmax = 30;                                    % maximum peak for Gaussian
[Phase, obj] = Single_Gaussian(Nx,Ny,Pmax);   % object/phase function
sigma = 0.0005;                               % width of Gaussian phase in m
```

```
fdelta = dx/sigma;
xx = ((1:Nx)-floor(2*Nx/4)).*fdelta;
xx = xx.*xx;
yy = xx;
[XX, YY] = meshgrid(xx,yy);
Int = 1;                                        % constant intensity distribution
obj = Int.*obj;
% figure of the designed phase
figure; imagesc(y_obj*1000,x_obj*1000,Phase); axis square; colormap
jet;colorbar;title('Object Phase');
xlabel('x (mm)');ylabel('y (mm)'); set(gca,'Fontsize',14);
xlim([min(x_obj*1000),max(x_obj*1000)]);ylim([min(y_obj*1000),max(y_obj*1000)]);



%% CGH Generation
thex = 1.5/180*pi;                  % tilt angle for off-axis geometry along x-axis
they = thex;                        % along y-axis
k0x = k0 * sin(thex);               % tilt angle introduced transverse spatial frequency
k0y = k0 * sin(they);
ref = exp(1i*(k0x*X + k0y.*Y));     % reference plane wave
for i = 1
H = transfer_fn(Nx,Ny,d,dx,dy,lambda1);          % non-paraxial transfer function
OBJ = fftshift(fft2(fftshift(obj)));             % spectrum of the object field
Complex_field = fftshift(ifft2(fftshift(OBJ.*H)));
Holo1 = abs(ref + Complex_field).^2;             % Hologram function (field)

% figure for the hologam
figure;
imagesc(y_obj*1000,x_obj*1000,Holo1);
colormap jet; axis square;
xlabel('x (mm)');ylabel('y (mm)'); set(gca,'Fontsize',14);
xlim([min(x_obj*1000),max(x_obj*1000)]);ylim([min(y_obj*1000),max(y_obj*1000)]);
title(['Holo',num2str(i)]);
```

```
axes('Position',[0.25 0.7 0.2 0.2]);imagesc(y_obj*1000,x_obj*1000,Holo1);
colormap jet; axis square;xlim([-2 -1.8]);ylim([-2 -1.8]);set(gca,'Fontsize',11);
end



%% Filtering
for i = 1
Holot1 =Holo1;
spec = fftshift(fft2(Holot1));                    % spectrum of the hologram

% find spectrum peak for first order frequency
fim_temp = abs(spec);
[xfim,yfim] = size(fim_temp);
siz = 10;
regx = ceil(xfim/2)-siz:ceil(xfim/2)+siz;
regy = ceil(yfim/2)-siz:ceil(yfim/2)+siz;
fim_temp(regx,regy) = 0;
fim_temp = fim_temp/max(max(fim_temp));
[xp,yp] = find(fim_temp == 1);
if length(xp) < 2
disp('Reference Error: cosine spectrum is not symetric')
end

% Design filter
fsize = 80;
xz = xp(2); yz = yp(2);
padx = (Nx - 2*fsize)/2;
pady = padx;
specf = spec(xz-fsize:xz+fsize-1,yz-fsize:yz+fsize-1);

% Filtering
specf = padarray(specf,[padx pady],0,'both');
obj_field1 = ifft2(fftshift(specf));
```

```
96
97  % figure of the filtered spectrum (absolute value)
98  figure;
99  surf(kx,ky,(abs(spec)));
100 colormap jet; colorbar; shading interp;
101 xlabel('kx (m^{-1})');ylabel('ky (m^{-1})'); set(gca,'Fontsize',14);
102 xlim([min(kx),max(ky)]);ylim([min(kx),max(ky)]);
103 end
104
105
106
107 %% Reconstruction
108 for i = 1
109 obj_temp = obj_field1;
110 H1 = inv_transfer_fn(Nx,Ny,-d-dzm(1),dx,dy,lambda1);    % inverse propagated to three
    planes
111 H2 = inv_transfer_fn(Nx,Ny,-d-dzm(2),dx,dy,lambda1);
112 H3 = inv_transfer_fn(Nx,Ny,-d-dzm(3),dx,dy,lambda1);
113 OBJ_temp = fftshift(fft2(fftshift(obj_temp)));
114 CF1 = fftshift(ifft2(fftshift(OBJ_temp.*H1)));
115 CF2 = fftshift(ifft2(fftshift(OBJ_temp.*H2)));
116 CF3 = fftshift(ifft2(fftshift(OBJ_temp.*H3)));
117
118 I1 = abs(CF1).^2;                                    % intensity of the obejct field at
    three planes
119 I2 = abs(CF2).^2;
120 I3 = abs(CF3).^2;
121
122 % figures of the intensity images
123 figure;
124 imagesc(y_obj*1000,x_obj*1000,I1);
125 colormap jet; colorbar; axis square;
126 xlabel('x (mm)');ylabel('y (mm)'); set(gca,'Fontsize',14);
127 xlim([min(x_obj*1000),max(x_obj*1000)]);ylim([min(y_obj*1000),max(y_obj*1000)]);
```

```
128
129 figure;
130 imagesc(y_obj*1000,x_obj*1000,I2);
131 colormap jet; colorbar; axis square;
132 xlabel('x (mm)');ylabel('y (mm)'); set(gca,'Fontsize',14);
133 xlim([min(x_obj*1000),max(x_obj*1000)]);ylim([min(y_obj*1000),max(y_obj*1000)]);
134
135 figure;
136 imagesc(y_obj*1000,x_obj*1000,I3);
137 colormap jet; colorbar; axis square;
138 xlabel('x (mm)');ylabel('y (mm)'); set(gca,'Fontsize',14);
139 xlim([min(x_obj*1000),max(x_obj*1000)]);ylim([min(y_obj*1000),max(y_obj*1000)]);
140
141 end
142
143
144
145 %% Solve TIE
146 tic
147 Step1 = k0.*(I3 - I1)./(2*1.*dz);
148 Step2 = fftshift(fft2((Step1)));
149 g = 1./(Kx.^2 + Ky.^2);
150 is = isinf(g);
151 g(is) = 0;
152 Step3 = g.*Step2;
153 Step4 = (ifft2(fftshift(Step3)));
154 Phase_TIE = real(Step4);
155 toc
156
157 % figure of the final result
158 figure;
159 surfl(x_obj*1000,y_obj*1000,Phase_TIE);colormap gray;shading interp;lighting flat;
160 set(gca, 'fontsize', 16, 'linewidth',1);view(37,47);
161 xlabel('x (mm)');ylabel('y (mm)');zlabel('z (rad)');
```

```
%% single Gaussian
function [Phase, obj] = Single_Gaussian(Nx,Ny,Pmax)
Nxx = Nx/2; Nyy = Ny/2;
sigma = 0.0005;                         % for generation of Gaussian Beam
delta = 4.65e-6;                        % pixel pitch in m
fdelta = delta/sigma;
Phase1 = zeros(Nx,Ny);

for i = 1:Nx
x = (i-floor(2*Nxx/2))*fdelta;
xx = x*x;
    for j = 1:Ny
        y = (j-floor(2*Nyy/2))*fdelta;
        yy = y*y;
        Phase1(i,j) = Pmax*exp(-(xx+yy));
    end
end

Phase = Phase1;

hreal = cos(Phase);                 % real part of the recorded complex amplitude
himag = sin(Phase);                 % imaginery part of the recorded complex amplitude
obj = complex(hreal,himag);    % complex amplitude in the plane of the phase object
% Int_obj = obj.*conj(obj);      % intensity of the object
end


%% Transfer function (non-paraxial) (forward propagation)
function H = transfer_fn(Nx,Ny,dzm,dx,dy,lambda)
fac1 = (2*pi/Nx)*dzm/dx;
```

```
fac2 = Nx*dx/lambda;
H = zeros(Nx,Ny);
for n = -Nx/2:Nx/2-1
   for m = -Ny/2:Ny/2-1
       argG = fac2^2-n*n-m*m;
       if argG <= 0.0
           H(n+Nx/2+1,m+Ny/2+1) = 0;
       else
           H(n+Nx/2+1,m+Ny/2+1)=exp(-1i*fac1*sqrt(argG));
       end
    end
end
end



%% Transfer function (non-paraxial) (backward propagation)
function H = inv_transfer_fn(Nx,Ny,dzm,dx,dy,lambda)
fac1 = (2*pi/Nx)*dzm/dx;
fac2 = Nx*dx/lambda;
H = zeros(Nx,Ny);
for n = -Nx/2:Nx/2-1
    for m = -Ny/2:Ny/2-1
       argG = fac2^2-n*n-m*m;
       if argG <= 0.0
           H(n+Nx/2+1,m+Ny/2+1) = 0;
       else
           H(n+Nx/2+1,m+Ny/2+1)=exp(1i*fac1*sqrt(argG));
        end
   end
end
end
```