Studies on resolution of digital holography system

Yan, Hao

2012

Yan, H. (2012). Studies on resolution of digital holography system. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/50764>

<https://doi.org/10.32657/10356/50764>

STUDIES ON RESOLUTION OF DIGITAL HOLOGRAPHY SYSTEM

SUBMITTED

BY

YAN HAO

Supervisor: Prof. Anand Asundi

**School of Mechanical and Aerospace Engineering
Nanyang Technological University**

Thesis submitted to Nanyang Technological University
In partial fulfillment of the requirements for the Degree of
Ph.D. (Mechanical & Aerospace Engineering)
Nanyang Technological University

2012

Abstract

Digital holography (DH) is an interferometry based quantitative phase measurement technique. It includes two processes. First a digital hologram is recorded using a CCD camera. Second numerical reconstruction is performed to determine the amplitude and phase. From the phase image, the 3D measurement of the object is possible.

In this thesis, the lateral resolution, its improvement and the axial measurement accuracy of digital holography are studied. Since lateral measurement and axial measurement are based on different mechanisms, digital holography system has different lateral and axial measurement capabilities. Therefore the lateral resolution and axial measurement accuracy are analyzed individually for lensless Fresnel holography configuration.

Firstly, the lateral resolution of lensless digital holography is limited by CCD specifications. Three factors contribute to this limitation, namely, pixel averaging effect within the finite detection size of one pixel, finite CCD aperture size limitation and sampling effect due to finite sampling interval. As DH system is space variant, influences of object extent on lateral resolution are also involved. Interactions of these factors on lateral resolution are investigated and presented. The lateral resolution of DH system can be determined for given parameters of these factors. The domains dominated by different factors are explained along with their accuracy. Lateral resolution performance of in-line and off-axis systems is also studied and examples of lateral resolution determination for a practical system are provided.

Secondly, with the results in the lateral resolution analysis, the improvement of lateral resolution is investigated by increasing the numerical aperture of the system with aperture synthesis method. Both the lateral resolution and image field of view can be enhanced at the same time using a more general Fresnel holography setup by hologram stitching. In the experiment, the synthesis is executed by moving the compact digital holographic system in two directions. Nine holograms are recorded and stitched into one hologram. The reconstruction results show that expanding aperture can improve lateral resolution.

In the last part, axial measurement errors of digital holography under the influences of different limitations are analyzed. The processes related with hologram recording are focused. Related factors are finite CCD size, pixel averaging due to the signal integral within single pixel detection size, sampling effect due to CCD camera and tilt angle between the reference and the object waves. The object placement also affects the system performance. The impacts of all the above factors on the axial measurement errors are analyzed. The influences of CCD size and object displacement on the axial accuracy are demonstrated with experiments.

Acknowledgment

I would like to express my sincere appreciation and gratitude to my supervisor Prof. Anand krishna Asundi for his introduction to the field of optical metrology and his invaluable guidance, patience, encouragement, constant trust, support in my PhD study.

I would like to thank Prof. Qian Kemao for his help in solving some key problems in my research work and his kind suggestions in building good research attitude and spirit.

I appreciate the patient guidance of Dr. Qu Weijuan in the beginning and her unreserved help during my PhD study. I give my thanks to Dr. Vijay Singh for advices and supports in the experiments and to Mr. Huang Lei and Zhu Hui for their research suggestions.

I am thankful to technicians Mr. Koh Hai Tong and Grace Ho and all the technicians of Micromachines, Mechanics of Materials, Metrology, Computer Aided Engineering, Precision Engineering, Manufacturing Process Labs for their help and support in work.

I would like to express my gratitude to my parents for their love as always and to my husband for his constant encouragement, support and comfort.

Finally I would like to acknowledge the financial support from NTU during my PhD study and the financial support from school MAE for international conference.

Declaration

I declare that this thesis has been composed by me and that the work contained herein is my own except where explicitly stated otherwise in the text. This work has not been submitted for any other degree or professional qualification.

However, some contents of this thesis had been published, accepted or submitted for publication of the following peer-reviewed journals and conferences:

Publications:

- **Hao Yan**, Anand Asundi, “Resolution analysis of a digital holography system”, APPLIED OPTICS, Vol. 50, Iss. 2, pp. 183–193 (2011).
- **Hao Yan**, Anand Asundi, "Studies on aperture synthesis in digital Fresnel holography," Optics and Lasers in Engineering, vol. 50, pp. 556-562.
- **Hao Yan**, Anand Asundi, “Comparison of Digital Holographic Microscope and Confocal Microscope methods for characterization of micro-optical diffractive components”, Proceedings of SPIE, Vol. 7155, Pages: B1550-B1550, 2008.
- **Hao Yan**, Ailing Tian, Anand Asundi, “Challenges of Digital Holography in Micro-optical Measurement” , Proceedings of SPIE, Vol. 7522, 2010.
- **Hao Yan**, Anand Asundi, “Calibration of Pizeo Rotation Nanopositioning Stage by Digital Holographic System”, Physics Procedia 19, 169-172 (2011).
- **Hao Yan**, Anand Asundi, "Resolution Analysis of In-line Digital Holography" in OSA Technical Digest (CD) (Optical Society of America, 2011), CWB4.

Table of Contents

Abstract.....	i
Acknowledgment.....	iii
Declaration	iv
Table of Contents	v
CHAPTER 1 INTRODUCTION.....	1
1.1 Introduction	1
1.2 Objective and Scope:.....	5
1.3 Chapter Organization.....	7
CHAPTER 2 LITERATURE REVIEW OF PHASE MEASUREMENT METHODS ...	9
2.1 Imaging Based Phase Measurement Techniques.....	10
2.1.1 Confocal Microscope.....	10
2.1.2 Quantitative Phase Microscope	14
2.2 Interferometry Based Phase Measurement Techniques.....	18
2.2.1 Phase Contrast Microscopy	18
2.2.2 Differential Interference Contrast (DIC) Microscopy	21
2.2.3 Fourier Phase Microscopy	25
2.2.4 Hilbert Phase Microscope.....	28
2.2.5 Diffraction Phase Microscope	32
2.2.6 Quantitative Differentiation Interference Contrast Microscope.....	34
2.2.7 Spectral-Domain Phase Microscope.....	37
2.2.8 Digital Holography	41
2.3 Conclusion.....	51
CHAPTER 3 DIGITAL HOLOGRAPHY	56

3.1	Configurations of DHM	57
3.2	Digital Recording	61
3.3	Numerical Reconstruction	66
3.3.1	Object Wavefield Reconstruction at Hologram Plane.....	66
3.3.2	Object Wavefront Propagation from Hologram Plane to Image Plane	70
3.4	Aberration Compensation.....	80
3.5	Review of Methods for Resolution Improvement in Digital Holography.....	86
3.5.1	Review of Lateral Resolution Analysis	86
3.5.2	Review of Lateral Resolution Improvement	90
3.5.3	Review of Axial Resolution	109
3.6	Possible Exploration Directions	111
CHAPTER 4 LATERAL RESOLUTION ANALYSIS OF DIGITAL HOLOGRAPHY		
		114
4.1	Introduction	114
4.2	Holography Expression with Finite CCD Size, Pixel Averaging and Sampling Effect	115
4.3	Finite CCD Size and Pixel Averaging Effect	117
4.3.1	Point Spread Function (PSF) Analysis of DH	117
4.3.2	Object with Finite Extent.....	125
4.4	Sampling Effect	134
4.5	Examples of System Analysis	134
4.5.1	In-line Geometry.....	134
4.5.2	Off-axis Geometry	135
4.6	Conclusion.....	135

CHAPTER 5 LATERAL RESOLUTION IMPROVEMENT BY APERTURE

SYNTHESIS.....	140
5.1 Introduction	140
5.2 Theoretical Analysis	141
5.2.1 Off-axis Fresnel DH Microscope	141
5.2.2 Aperture Synthesis.....	145
5.3 Experiment and Results	150
5.3.1 Hologram Stitching	150
5.3.2 Lateral Resolution and FOV Improvements.....	152
5.4 Conclusion.....	154
CHAPTER 6 ANALYSIS OF AXIAL MEASUREMENT ERRORS.....	158
6.1 Introduction	158
6.2 PSF of DH System.....	160
6.3 Ideal Cases.....	173
6.3.1 Infinitely large CCD Size and Infinitely Small Pixel Size	173
6.3.2 Infinitely Small Pixel Size.....	174
6.3.3 Infinitely Large CCD Size.....	174
6.3.4 Summary.....	177
6.4 Sampling Effect.....	179
6.5 Investigation of Axial Measurement Errors of Object with Extent.....	179
6.5.1 Influence of Finite CCD Size	181
6.5.2 Influence of Object Position	183
6.5.3 Influence of Carrier Frequency	185
6.5.4 Influence of Pixel Size.....	187
6.6 Experiment and Results	189

6.6.1	Finite CCD Size.....	189
6.6.2	Object Displacement.....	197
6.7	Conclusion.....	200
CHAPTER 7 CONCLUSION AND FUTUER WORK.....		202
7.1	Conclusion.....	202
7.2	Future Works	205
REFERENCES		207

CHAPTER 1 INTRODUCTION

1.1 Introduction

In digital holography, the object wavefront is coded in the hologram during the recording process through interference with a reference wavefront. The object wavefront is then retrieved by numerical reconstruction of hologram. From the digital object wavefront, its amplitude and the phase can be extracted. The amplitude provides the 2D intensity image while the phase provides the optical path length difference (OPD).

The phase is proportional to the OPD as in equation $Phase = 2\pi \times \frac{OPD}{\lambda}$. Assuming a plane reference wavefront, the OPD is directly related to the object wave. When optical wave travels through a transmitted specimen, the change in amplitude is due to the absorption ability of specimen and the change in phase is related to the optical thickness of specimen. The relationship between phase and a transmitted specimen is presented in Fig. 1.1 (a) and (b). The phase map obtained from the object wavefront at the observation plane actually presents specimen thickness distribution with known reference index n of the object in transmission mode. When optical wave is reflected by a reflective specimen, the change in amplitude is due to the reflectivity of specimen and the change in phase is related to the profile of specimen (Phase shift of half wavelength exist in reflection. But it does not affect the measurement of specimen profile). The relationship between phase and a reflective object is presented in Fig. 1.1 (c) and (d).

CHAPTER 1 INTRODUCTION

The phase map obtained from the object wavefront at the observation plane actually presents specimen profile distribution in reflection mode.

Figure 1.1 Relationship between phase and specimen thickness in transmission mode and specimen height in reflection mode. (a) Transmitted object with light travelling through it; (b) phase map of transmitted

CHAPTER 1 INTRODUCTION

object at the observation plane; (c) reflective object with light illuminated from top of it; (d) phase map of reflective object at the observation plane.

As the lateral and axial dimensional measurements are based on different mechanisms, digital holography system has different lateral resolution and axial resolution capabilities. The lateral resolution is limited by the diffraction effect while the axial resolution is not. Due to the diffraction limit, the lateral resolution cannot go beyond sub-micrometer level. The axial measurement is based on OPD measurement of the object wavefront, it is said to have of several nanometres resolution capability. Therefore the lateral resolution and axial resolution should be analyzed individually.

The definition of lateral resolution is used to describe the capability of a system to laterally distinguish two nearby points. Rayleigh criterion is the most often used criteria to define the lateral resolution of an imaging system. But it may not be appropriate to define the lateral resolution of digital holographic system in certain aspects. First, Rayleigh criterion is based on incoherent system while digital holography is a coherent system. In a coherent system, whether two nearby points with lateral distance indicated by Rayleigh criterion can be resolved not only depends on their lateral distance but also depends on their relative phase difference. Second, Rayleigh criterion is derived for an imaging system while digital holography is not a pure imaging system. It is interferometry in which the CCD recording and the numerical calculation of wave propagation are involved in the image reconstruction. In this case not only the aperture but also other system factors related to interferometry, digital recording and numerical reconstruction may affect the lateral resolution. Therefore the lateral resolution of digital holography needs further analysis and definition.

CHAPTER 1 INTRODUCTION

Digital holography is a three dimensional measurement technique. Like the lateral resolution, the axial resolution is also a very important parameter to define the system performance in axial measurement. But there is no such criterion to define the axial resolution as Rayleigh criterion for the lateral resolution. The concept of depth resolution $\frac{\lambda}{2(NA)^2}$ has been addressed which describes the depth distance at which the intensity of two points can be resolved. But it is not equivalent to axial resolution. The depth resolution is based on the 2D intensity image while the axial resolution is based on the 3D phase image. They are derived from two different mechanisms and therefore are two different concepts. Furthermore the depth resolution $\frac{\lambda}{2(NA)^2}$ is derived from the imaging system as the Rayleigh criterion. Whether the equation $\frac{\lambda}{2(NA)^2}$ is sufficient and eligible to define the depth resolution of digital holography still needs justification.

The quantization effect due to the camera gray scale levels in the process of analog-to-digital conversion (ADC) of the hologram has been considered. As the OPD information represented in the phase of the object wave is encoded by the reference wave into the interference pattern--hologram. Therefore the quantization involving rounding or truncating of the intensity values of hologram causes phase and therefore OPD errors. If the ADC digitizes analog signal into an eight bit data, there is $2^8=256$ discrete quantization levels. The quantization effect gives phase resolution of $\frac{2\pi}{256}$ which corresponds to OPD resolution of $\frac{\lambda}{256}$. This OPD resolution provides an axial resolution of $\frac{\lambda}{256 \times (n-1)}$ in transmission mode and an axial resolution of $\frac{\lambda}{256 \times 2}$ in reflection mode.

However, such axial resolution cannot be achieved in practice which has been manifested by reported works. One of the reasons may be the relatively large axial measurement errors compared to the axial resolution predicated by theory. In order to achieve high axial resolution, axial measurement errors need to be identified first and then be compensated.

In this thesis, the investigations of lateral resolution and axial measurement errors are performed individually. The influences of different practical systematic factors on the lateral resolution and the axial measurement errors are identified and analyzed. The mechanisms of their interactions on the lateral resolution and the axial measurement errors are analyzed. The ways to improve lateral resolutions are investigated and implemented.

1.2 Objective and Scope:

Objective of this study is to analyze and enhance the lateral resolution and axial measurement accuracy of lensless digital holography using Fresnel geometry.

The scope of the work is:

1) Lateral Resolution Analysis of Digital Holography System

The lateral resolution of digital holography is limited due to CCD or other recording devices. Three factors contribute to this limitation, namely, pixel averaging effect within the finite detection size of one pixel, finite CCD aperture size and sampling effect due to the finite sampling interval. In this part, interactions of the three factors on lateral

resolution are investigated and presented. The lateral resolution of DH system can be determined for given parameters of these three factors. The domains dominated by different factors are explained along with their accuracy. As DH system is space variant, influences of object extent on the lateral resolution are also discussed. The lateral resolution performance of in-line and off-axis systems is studied and examples of the lateral resolution determination for a practical system are provided.

2) Lateral Resolution Improvement

With the results in the work of lateral resolution analysis, the improvement of lateral resolution is investigated by increasing the numerical aperture of the system with aperture synthesis method. In this part of work, both the lateral resolution and image field of view are enhanced at the same time using a more general Fresnel holography setup and hologram stitching. The impact of aperture synthesis on the lateral resolution is investigated theoretically and experimentally. In the experiment, the synthesis is executed by moving the compact digital holographic system in two directions. Nine holograms are recorded and stitched into one hologram. The reconstruction results show that expanding aperture can improve the lateral resolution.

3) Analysis of Axial Measurement Accuracy of Digital Holography System

In this part of work, axial measurement accuracy of digital holography under the influences of different limitations is analyzed. The systematic processes related to the hologram recording are focused. Factors including finite CCD size, pixel averaging due to the integral of signal within single pixel detection size, sampling effect due to CCD camera are discussed. The reference wave to produce hologram at the CCD and its

conjugate which extracts wavefront at the CCD from hologram are also included. As DH system is space variant, object placement also affects the system performance. Their impacts on the axial measurement errors are analyzed. Experiments on the influences of CCD size and object displacement agree with the analysis of these two factors on the axial accuracy.

1.3 Chapter Organization

Organization of this thesis is as follows:

In chapter 2, different phase imaging and measurement techniques are reviewed and the principles and characteristics of these techniques are discussed. Their advantages and disadvantages are compared. The choice of digital holographic microscope is discussed.

In chapter 3, the basic theory of digital holography and different configurations of holography are introduced. Procedures and principles of digital hologram recording and numerical optical field reconstruction are presented. We provide a review of compensation methods of curvatures due to microscope objective (MO), different curvatures of reference and object waves and etc. Reviews of studies on the lateral resolution, its improvement and axial resolution and accuracy of digital holography are given. Based on the reviews, possible exploration directions are pointed out at the end.

In chapter 4, the interactions of the systematic limitations on the lateral resolution are investigated and presented. System parameters involved are pixel averaging effect within the finite detection size of one pixel, finite CCD aperture size, sampling effect

CHAPTER 1 INTRODUCTION

due to finite sampling interval and object extent. Lateral resolution performance for a practical system is provided.

In chapter 5, with the results in the work of lateral resolution analysis, the improvement of lateral resolution is achieved by increasing the numerical aperture of the system with aperture synthesis method. Both the lateral resolution and image field of view can be enhanced at the same time using a more general Fresnel holography setup.

In chapter 6, the axial measurement accuracy of digital holography under the influences of different limitations is analyzed. The focused factors are finite CCD size, pixel averaging due to the integral of signal within single pixel detection size, sampling effect due to CCD camera, carrier frequency introduced by reference wave and the space variant property of digital holography. Experiments on the influences of CCD size and object displacement agree with the analysis of these two factors on the axial accuracy. The conclusions achieved in this chapter can be used as a guide for DH system adjustment to improve the axial measurement accuracy.

In chapter 7, contributions of this thesis are concluded. Strengths and weaknesses of current work are summarized. Future works are discussed.

CHAPTER 2 LITERATURE REVIEW OF PHASE MEASUREMENT METHODS

In this section, a general view of some important phase measurement techniques is provided.

Phase measurement technique differs from intensity measurement technique such as conventional microscope in that it can provide 3D information including the additional axial dimension information instead of only the 2D intensity information. This property makes it obviously advantageous in the visualization and measurement, especially for phase specimens. When light propagates through a phase specimen, light is hardly absorbed by it and light intensity will not be changed obviously. Therefore phase specimen is mostly transparent and is problematic for conventional microscopy which is based on intensity observation. Many microscopic biological specimens, such as cells and their intracellular constituents, and microlens are phase specimens. As the phase measurement technique detects the phase of the light instead of the intensity of the light and the phase are related to the optical thickness difference of phase specimen, it opens up another way to visualize and measure the phase specimen.

There are a number of known techniques to qualitatively or quantitatively measure the direct phase changes or various indirect phase related information. The mechanisms of these phase measurement techniques can be divided into two categories: imaging related techniques and interferometry related techniques. The techniques in this chapter are

organized in these two categories. According to the final result, the techniques can be characterized as direct phase measurement technique and indirect but phase related measurement technique. According to the result format, the techniques can be characterized as quantitative and qualitative phase measurement techniques.

2.1 Imaging Based Phase Measurement Techniques

2.1.1 Confocal Microscope

The confocal microscope [1] is an established technique for three-dimensional microscopic imaging in areas ranging from biology and medicine to industrial micro-inspection. Figure 2.1 shows basic geometrical structure of confocal microscope, based on the principle of optical sectioning. As seen in Fig. 2.1 (a), a laser beam passes through a pinhole aperture and then focused by an objective lens in to a small focus volume within the specimen. The reflected laser light from the illuminated spot is then recollected by the objective lens again. After passing through the beam splitter, the light arrives at another pinhole aperture at the detector. This aperture obstructs the light that is not coming from the focal plane, as shown by the dotted line in Fig. 2.1 (a). A detector is placed behind this pinhole aperture to detect the intensity of light passing through it. If the intensity is higher than the threshold, it is considered as the intensity from focal point and is recorded. If the intensity is lower than the threshold, the light is considered to come from out-of-focus plane and the intensity will not be recorded. The relationship between the intensity at the detector and distance between the plane where light comes from and the focal plane is shown in Fig. 2.1 (b). As most of the returning

light is blocked by the pinhole, the out-focus-image is suppressed, resulting in sharper images than those from the conventional microscopes. The pinhole plane at the laser source and the pinhole plane at the detector are confocal planes with each other. Thus such microscope is named as confocal microscope. The detected light originating from an illuminated volume within the specimen represents a pixel in the resulting image. As laser scans pixel by pixel, line by line, a 2D image of the specimen can be generated. If laser continues to scan plane by plane, a 3D image of the specimen can be generated.

(a)

(b)

CHAPTER 2 LITERATURE REVIEW OF PHASE MEASUREMENT METHODS

Figure 2.1 (a) Sketch map of confocal microscope [2]; (b) Relationship between the light intensity at the detector and distance between the plane where light comes from and focal plane.

The confocal method is particularly valuable in 2D fluorescence microscopy [3] in which the illuminating beam excites fluorescence from parts of the sample. Figure 2.2 is a typical set of pictures taken by conventional bright-field and fluorescence confocal microscope respectively. Figure 2.2 (a), (c) and (e) are taken by conventional microscopy while (b), (d) and (f) are taken by fluorescence confocal microscope. a) and (b) are Mouse brain hippocampus section; (c) and (d) are Rat smooth muscle; (e) and (d) are Sunflower pollen grain. It can be seen that the images of confocal microscopy are much clearer than those of conventional microscopy.

Figure 2.2 (a) and (b) are images taken from Mouse brain hippocampus section; (c) and (d) are images taken from Rat smooth muscle; (e) and (d) are images taken from Sunflower pollen grain; (a), (c) and (e) are taken by conventional microscopy; (b), (d) and (f) are taken by fluorescence confocal microscope [1].

Another important application of confocal microscopy is in 3D imaging by optical sectioning. In this kind of confocal microscopy, the confocal pinhole planes are moved to take 2D confocal image of different layers in specimen and at the same time record their layer positions. 3D structure can be then generated with the help of deconvolution software which combines all the 2D images according to their positions. The thickness of this optical section can be sub-micron if an objective of high numerical aperture is used, giving sub-micron axial resolution. Figure 2.3 shows a 3D image and profile of a micro-optics taken by reflection confocal microscope of MAE school in NTU.

For phase object, when light propagates through them, phase variation is proportional to OPD which is related to the thickness of the object and the refractive index of the specimen. For those phase objects with constant refractive index, such as the transparent micro-optics in Fig. 2.3 (a), phase change is determined only by the thickness of specimen. As confocal microscope could provide quantitative thickness information of such phase object, it can be used to supply quantitative phase information for such phase objects with constant refractive index. Therefore confocal microscope is introduced as a phase measurement technique in this chapter.

Figure 2.3 (a) 3D image of a micro-optics; (b) Profile along center line through the sample.

For fluorescence confocal microscope, the organelles or substances in cells which can react with fluorescence can be seen with high contrast. Specimens need to be dyed with fluorescence. And only qualitative 2D intensity image is provided.

For 3D profiling confocal microscope, it can provide quantitative and direct phase information for those phase objects with constant refractive index. However, the speed is limited by scanning. And for it measures height variation, it can provide quantitative phase information only for those phase objects with constant refractive index. Also the sectioning depth is determined by the depth of field. There is limitation for sectioning depth—it is difficult to be less than 0.2 micrometer. Thus specimens with height less than 0.2 micrometer and features smaller than 0.2 micrometer could not be detected by confocal microscope.

2.1.2 Quantitative Phase Microscope

Quantitative phase microscope technique[4] is based on the Transport of Intensity Equation (TIE) proposed by Teague [5, 6]:

$$\frac{2\pi}{\lambda} \frac{\partial}{\partial z} I(x, y, z) = -\nabla_{xy} \cdot [I(x, y, z) \nabla_{xy} \phi(x, y, z)] \quad (2.1)$$

with wavefield of $\psi(x, y, z) = \sqrt{I(x, y, z)} \exp \{i\phi(x, y, z)\}$. ∇_{xy} is a derivative operator working in the x - y plane. This equation presents the relationship between phase $\phi(x, y, z)$ and intensity $I(x, y, z)$ in case of paraxial wave propagation. With known intensity $I(x, y, z)$ and its differential $\frac{\partial}{\partial z} I(x, y, z)$, phase $\phi(x, y, z)$ can be determined by solving Eq. (2.1).

Different solutions of TIE have been proposed such as Poisson auxiliary function[5], Fourier transform implementation[7], iterative approach[8], etc. And performance comparisons of different approaches have been discussed[8].

Normally this technique utilizes conventional bright-field transmission microscope to collect the intensity of in-focus image $I(x, y, z)$ and very slightly positively and negatively defocused images[6, 9] as shown in Fig. 2.4 and uses these intensity data to estimate the intensity differential $\frac{\partial}{\partial z} I(x, y, z)$ in the middle plane as shown below:

$$\frac{\partial}{\partial z} I(x, y, z) = \frac{I(x, y, z+\Delta z) - I(x, y, z-\Delta z)}{2\Delta z} \quad (2.2)$$

The phase distribution $\phi(x, y, z)$ at the middle plane is then determined from image intensities at three image planes by solving Eq. (2.1).

Figure 2.4 Three images taken with the same defocus step [6].

Applications of quantitative assessment of cell attributes have been developed [10] such as tracking of culture confluency and growth to investigate cell proliferative properties [11], cell volume [12] and cell refractive index measurement [13]. Other phase objects, such as optical fiber, have been investigated by this technique as well [13].

An example of application is shown in Fig. 2.5 (a) and (c) are bright-field image of human buccal epithelial cell and mouse erythrocyte respectively. (b) and (d) are the phase images retrieved by quantitative phase microscopy from (a) and (c) respectively.

CHAPTER 2 LITERATURE REVIEW OF PHASE MEASUREMENT METHODS

Figure 2.5 (a) Bright-field image of human buccal epithelial cell (Achromplan, X40; numerical aperture (NA), 0.60). (b) Phase map of the cell shown in (a), showing a prominent phase dense (darker) nucleus. (c) Bright-field image of mouse erythrocyte (Achromplan, X63, NA, 0.80). (d) Phase map of the erythrocyte shown in (c), with biconcavity depicted as a darkened annulus of increased phase [14].

As the name suggests quantitative phase microscopy provides quantitative direct phase information. As it is based on conventional transmission microscopy, it is an imaging based phase measurement technique. It can only measure transparent objects especially phase object like unstained live cells and provide phase information $(n_1 - n_2) \times d \times \frac{2\pi}{\lambda}$ where n_1 is refractive index of cell, n_2 is the refractive index of environment where cell is in and d is thickness of specimen. It cannot measure phase based on reflection geometry. Its recording content includes one in-focused intensity image and two or more defocused intensity images. The final calculated result is quantitative phase and intensity distribution at the in-focused plane. The reconstruction method to acquire the final result from the recording content is solving the Transport Intensity Equation. This technique is a non-interference approach and no specimen preparation is needed.

Advantages of quantitative phase microscope are that it is a non-destructive quantitative phase measurement technique; it is optically and practically simple, requiring only conventional transmission microscope and CCD camera; it is a non-interference technique so that the phase does not have to be unwrapped.

Shortcomings also exist. User needs to displace the sample through the focus and collect at least three intensity images at three different positions. This may limit its

applicability to dynamic biological studies; Reconstruction methods need extensive computation; it can only be applied to transparent specimen and is based on a transmission microscope.

2.2 Interferometry Based Phase Measurement Techniques

2.2.1 Phase Contrast Microscopy

Phase contrast microscope was invented by Zernike [15-17] for which he received the Nobel Prize in physics in 1953. This technique is now a widely utilized and effective microscopy technique to observe phase objects. The principle of it is translation of minute variations in phase into corresponding changes in amplitude which can be visualized as differences in image amplitude contrast. As this technique obtains phase information by converting the phase changes into observable amplitude variations, it is an indirect phase measurement technique. But the obtained amplitude image cannot be used to derive the according phase value. Hence this technique is a qualitative phase measurement technique.

A beam of light passing through a phase specimen as shown in Fig. 2.6 is divided into two components: One is the undiffracted wave (S wave) which is the primary component and passes through and around the specimen but does not interact with it; the other is diffracted spherical wave (D-wave) which has the components scattered in many directions as shown in Fig. 2.6.

Figure 2.6 Interaction of light with phase specimen[3].

Phase relationships between S wave and D wave are shown in Fig. 2.7. P-wave ($P = S + D$) is the interference of S and D. The D wave has lower amplitude and is retarded in phase of approximately 90° ($\lambda/4$) relative to S wave by the specimen. Thus the amplitudes of the S and P waves are nearly the same and the transparent specimen completely lacks contrast and is almost invisible compared to the bright background.

Figure 2.7 Phase relationships between S, D, and P waves in bright field condition [3].

Figure 4

Figure 2.8 Geometrical sketch map of phase contrast microscope [3].

The geometrical sketch map of phase contrast microscope is presented in Fig. 2.8. By positioning a condenser annulus in the front focal plane of condenser, D waves and S wave are segregated at the objective rear focal plane. By inserting a phase plate at the objective rear focal plane, the amplitude of S wave is reduced to be nearly equal to D wave and the phase of S wave is advanced or retarded (we take advance as example) by $\lambda/4$.

Figure 2.9 (1) and (2) Relationship between D and S, P and S wave after passing through phase plate respectively.

Hence the recombination of D and S waves is almost a destructive interference due to a 180° ($\lambda/2$) phase shifted D-wave as shown in Fig. 2.9. The 180° phase shift results from

90° retardation of D-wave introduced by the phase object and 90° phase advancement by S-wave due to the phase plate. Thus Phase specimen seems darker comparing with the background as shown in Fig. 2.10 (a). If phase plate retards the phase of S wave by 90° ($\lambda/4$), Phase specimen seems brighter comparing with the background as shown in Fig. 2.10 (b).

Figure 2.10 (a) human blood cell under positive phase contrast microscope with 100X objective lens ; (b) Human red blood cell under negative phase contrast microscope [3].

With phase contrast illumination "invisible" phase variations are hence translated into visible amplitude variations. This makes phase microscope one of the most widely used phase imaging technique. Halo and shade-off artifacts are shortcomings of phase microscope. It is also not suitable to measure phase objects with phase shifts larger than $\pi/2$. Its critical drawback is that it could not provide quantitative phase information, though it can make phase objects visible.

2.2.2 Differential Interference Contrast (DIC) Microscopy

Differential interference contrast microscopy[18, 19] (DIC) is an optical microscopy illumination technique used to enhance the contrast in unstained, transparent samples.

The principle of DIC is based on interferometry to gain information of the optical density of the sample, to see otherwise invisible features.

DIC works by separating a polarized light source into two beams which take slightly different paths through the sample. Interference of the two beams with different optical path length gives the appearance of a three-dimensional physical relief due to the variation of optical density of the sample, emphasizing lines and edges though not providing an accurate topographical image. Both the DIC and phase contrast microscopes are commonly used commercial microscopes for the phase objects imaging. However, the same as phase contrast microscopes, the amplitude distributions do not quantitatively map the phase distribution. They are qualitative method. Furthermore, these techniques entangle the phase information into amplitude images and observe incomplete phase information from the amplitude images. Therefore this technique is an indirect phase measurement technique.

Figure 2.11 Schematic of DIC microscope [20].

As shown the schematic of DIC microscope in Fig. 2.11, unpolarized light enters the microscope and is polarized at 45° at the polarizing filter. The polarized light then enters the first Nomarski-modified Wollaston prism and is separated into two rays polarized at 90° to each other, the sampling and reference rays. The two rays are focused by the condenser so that they will pass through two adjacent points in the sample, around $0.2\mu\text{m}$ apart. Then the rays travel through the different but adjacent areas of the sample. Thus in the sample they will experience different optical path lengths due to difference refractive index or thickness in different areas. This causes a change in phase of one ray relative to the other due to the delay experienced by the wave in the more optically dense material. These rays of different phase delays travel through the objective lens and are focused at the second Wollaston prism. The two rays are recombined into one beam polarized at 135° . This combination of the rays leads to interference, brightening or darkening the image at that point according to the optical path difference.

Figure 2.12 Example of DIC microscope: (a) A transparent specimen[21]; (b) DIC image of specimen [21]; (c) Intensity along the dotted line in (b).

Figure 2.12 shows an example of DIC microscope image. (a) is the transparent specimen and (b) is its DIC image which is the interference result of the two images generated by two perpendicularly polarized beams.(c) is the intensity along the dotted line of (b). The phase difference becomes visible through interference and this clearly shows the shape of the transparent sample.

DIC has strong advantages in imaging of live and unstained biological samples. Its resolution and clarity are unrivaled among standard optical microscopy techniques. The

main limitation of DIC is its requirement for a transparent sample of similar refractive index to its surroundings. DIC is unsuitable (in biology) for thick samples, such as tissue slices, and highly pigmented cells. DIC is also unsuitable for most non biological uses due to its dependence on polarization, which many physical samples may affect. Furthermore, analysis of DIC images must take into account the orientation of the Wollaston prisms and direction of illuminating light.

2.2.3 Fourier Phase Microscopy

In Fourier phase microscopy (FPM), the optical field associated with a microscope image is decomposed into high spatial frequency (ac) component and an average field (dc) in Fourier domain [10]. Phase-shifting interferometry technique is utilized to retrieve quantitatively the phase of the sample field. It provides the direct phase map.

Setup of Fourier phase microscopy is shown in Fig. 2.13. Details of it are described in reference [22]. Transmission microscope is used to generate magnified image wavefield at IP_1 . The image is then transferred to a CCD using a $4f$ system composed of lenses L2 and L4. The Fourier lens L2 spatially decomposes the image field into average dc component and a spatial varying or scattered field ac component at FP_1 plane. A phase contrast filter (PCF) is placed at the Fourier plane (FP_1) to generate phase shifts for phase-shifting interferometry algorithm. The center of the PCF is removed so that only the outer part of PCF can be phase modulated. The position of the dc component is adjusted to overlap the central pinhole of PCF such that only ac component undergoes

phase shifting of controlled value. The dc and phase-shifted ac components interfere at the image plane IP_2 of the 4f system. Interference pattern is captured by CCD.

Figure 2.13 Experiment set-up of Fourier phase microscopy [23].

Relative phase difference $\Delta\phi$ between the ac and dc components is acquired by four-frame phase-shifting interferometry algorithm. The intensity image recorded by CCD as a function of the phase shift increment has the form of Eq. (2.3) [23]:

$$I(x, y; n) = |E_0|^2 + |E_1(x, y)|^2 + 2|E_0||E_1| \times \cos[\Delta\phi(x, y) + n\pi / 2], \quad n=0, 1, 2, 3. \quad (2.3)$$

where $\Delta\phi$ represents the phase difference between E_0 (dc) and E_1 (ac) and is retrieved from the four interferograms by means of $\tan(\Delta\phi) = [I(3) - I(1)] / [I(0) - I(2)]$.

Thus the phase of the microscope image (the complex sum $E_0 + E_1(x, y)$) which is the quantity of interest, can be expressed as

$$\phi(x, y) = \tan^{-1} \left\{ \frac{\beta(x, y) \sin[\Delta\phi(x, y)]}{1 + \beta(x, y) \cos[\Delta\phi(x, y)]} \right\} \quad (2.4)$$

with $\beta(x, y) = |E_1(x, y)| / |E_0|$. Since E_0 is a plane wave, β can be expressed as

$$\beta(x, y) = \gamma \frac{[I(x, y; 0) - I(x, y; 2) + I(x, y; 3) - I(x, y; 1)]}{\sin[\Delta\phi(x, y)] + \cos[\Delta\phi(x, y)]} \quad (2.5)$$

The quantity $\phi(x, y)$ is therefore uniquely determined from the four interferograms with no additional measurements or inherent experimental complications.

Fourier phase microscopy has been applied to image live cell in culture [22]. Specimen preparation is not required. An example of application is shown in Fig. 2.14. (a) is Fourier phase microscope image of a live HeLa cell and (b) is digital DIC image of the same cell.

Figure 2.14 (a) Fourier phase microscope image of a live HeLa cell; (b) Digital DIC image obtained from the image in (a) [22, 23].

Fourier phase microscopy is a quantitative phase measurement technique based on interferometry. It can be interfaced with an existing conventional optical transmission microscope. This technique is stable as it utilizes the unscattered light transmitted through the sample as a reference for a common-path interferometer. It can extract quantitative phase images with sub-nanometer path-length sensitivity overtime period from seconds to a cell life cycle due to common-path geometry. The disadvantage of this technique is that it needs to record four interferograms for one phase image. And the setup is complicate and not easy to adjust and use. More optics involves more aberrations and limitations.

2.2.4 Hilbert Phase Microscope

Hilbert phase microscope (HPM) [23] is an optical interference based technique for quantitative phase imaging by retrieving a full-field phase image from a single spatial interferogram recording.

Typical set-up of Hilbert phase microscope is shown in Fig. 2.15. Details of this set-up are described in reference [24]. A laser beam is divided into two parts. One part in sample arm serves as the illumination field for the inverted microscope. A tube lens images the sample on the CCD via beam splitter cube. The other part of laser beam in reference arm is collimated and expanded by a telescopic system consisting of another microscope objective and the tube lens. This planar reference field interferes with image field with designed tilt angle to produce uniform fringes of an angle of 45° with respect to x and y axes. The intensity of recorded interferogram in one direction is in the form of Eq. (2.6) [25]:

$$I(x) = I_R + I_s(x) + 2\sqrt{I_R I_s(x)} \cos[qx + \phi(x)] \quad (2.6)$$

where I_R and I_s are, respectively, the reference and sample intensity distribution, q is the spatial frequency of the fringes, and ϕ is the spatial varying phase associated with objects which is aimed to measure.

Figure 2.15 Typical Hilbert phase microscope set-up [24].

First the interferogram is Fourier transformed and high-pass filtered to obtain sinusoidal term $u(x) = 2\sqrt{I_R I_s(x)} \cos[qx + \phi(x)]$. Then complex analytical signal is constructed as follows:

$$z(x) = \frac{1}{2}u(x) + i \frac{P}{2\pi} \int_{-\infty}^{\infty} \frac{u(x')}{x - x'} dx' \quad (2.7)$$

The imaginary part of the right-hand side is the Hilbert transform of $u(x)$. The relationship below exists according to the properties of Hilbert transform:

$$\phi(x) = \tan^{-1} \{ \text{Im}[z(x)] / \text{Re}[z(x)] \} - qx \quad (2.8)$$

Thus phase associated with object is acquired.

The process of computation of Eq. (2.7) to obtain complex analytic signal is equivalent to performing Fourier transform of two-dimensional sinusoidal signal and suppressing the negative spatial frequencies. From an inverse Fourier transform operation, a two-dimensional complex analytic signal could be obtained. And phase information can be deduced from it.

Figure 2.16 (a) and (c) are Hilbert phase microscope images for quantitative assessment of shape transformation of a red blood cell in a 10s period. (b) and (d) are measured along the profiles indicated by the arrows in (a) and (c) [26].

Hilbert phase microscope has been applied to retrieve phase profile of an optical fibre [25], quantify cell volume and monitor cell dynamic morphology at the millisecond scales and sub-nanometre path-length sensitivity [24], quantify the refractive properties

of pathology tissue slices [25]. An example of applications is shown in Fig. 2.16. (a) and (c) are images of Hilbert phase microscope for quantitative assessment of shape transformation of a red blood cell in a 10s period. (b) and (d) are measured along the profiles indicated by the arrows in (a) and (c).

The Hilbert phase microscope is a quantitative phase measurement approach based on interference technique and transmission geometry. What is recorded by such technique is the interferogram of image field with the planar reference field. The reconstruction method is isolation of the phase associating the object from the phase of a complex analytical signal constructed with help of Hilbert transform. The final result is quantitative phase or quantitative OPD which is optical thickness related to n and d where n is refractive index and d is the physical thickness of specimen. No specimen preparation is needed.

The Hilbert phase microscope can be used to accurately quantify nanometre-level path-length shifts at the millisecond time scales or less. This is due to its single shot nature so that its acquisition time is limited only by the recording device.

However, the optical system of Hilbert phase microscope is complex. Thus it is difficult to adjust and the phase aberration due to optical elements in the system will be complicate. Furthermore this technique can only reconstruct the phase information at image plane instead of volume as the interferogram is recorded at the image field.

2.2.5 Diffraction Phase Microscope

Diffraction phase microscope (DPM) [25] combines the single shot benefit of Hilbert phase microscope with the common path geometry of Fourier phase microscope. Thus, DPM allows for fast imaging rates without compromising phase stability.

The basic setup of the DPM is shown in Fig. 2.17. A grating is placed at image plane of specimen, which generates multiple diffraction orders containing full spatial information of the sample image. L3 and L4 and a pinhole form a 4-f spatial filtering system. This system isolates the 0th order and 1st order to generate a common path Mach-Zender interferometer. 0th order is the reference beam and 1st order is the object beam and they interfere at CCD plane. From the recorded interferogram, Hilbert transform is used to extract quantitative phase image as in Hilbert phase microscope [27]. Details of this setup are introduced in reference [24].

Figure 2.17 Diffraction phase microscope setup [28].

DPM technique has been combined with fluorescence microscopy [29] and confocal microscopy [30] techniques to enhance its capability.

DPM has been applied to quantitatively assess the single red blood cell shape and dynamics[31], monitoring cell attack phenomenon [27], red blood cell membrane fluctuation [31], particle tracing [29]. One of the applications is shown in Fig. 2.18. It is the profile of red blood cell. Since red blood cell has uniform refractive index, the phase image $\Delta\phi(x,y)$ is directly proportion to the height profile $h(x,y)$ where k is a constant. The scale bar on the right shows cell thickness in microns.

Figure 2.18 Quantitative phase image of a red blood cell of DPM. The scale bar on the right shows cell thickness in microns [29].

Diffraction phase microscope is a quantitative phase measurement approach based on interference technique and transmission geometry. What is recorded by such technique

is the interferogram of the image field with the planar reference field. The reconstruction method is isolation of the phase associating the object from the phase of a complex analytical signal constructed with help of Hilbert transform. The final result is quantitative phase or quantitative OPD which is optical thickness related to n and d where n is refractive index and d is the physical thickness of specimen. No specimen preparation is needed.

Diffraction phase microscope can be used to accurately quantify nanometer-level path-length shifts in millisecond time with sub-nanometer path-length stability. This is due to its single shot nature and common-path interferometer geometry.

However, many optics components such as several lenses, grating and pinholes are used in diffraction phase microscope. This makes the setup of DPM complicate and not easy to adjust and use. Phase aberration due to different optics components will be complicate and severe. Only the phase information at image plane can be reconstructed as the interferogram is recorded at the image field.

2.2.6 Quantitative Differentiation Interference Contrast Microscope

DIC microscope is one of the widely used phase imaging technique which is able to capture minute structures of phase objects as we discussed above. However, it is designed for qualitative phase imaging only. The image of DIC microscope is a mixture of amplitude information and phase gradient information wrapped in sinusoidal signal. But several approaches to achieve quantitative DIC have been proposed and reported.

Quantitative phase information can be extracted from DIC image— an entanglement of amplitude with phase gradient.

The first step for quantitative DIC is to extract phase gradient from DIC image. Phase-shifting method [29] and approximation methods [32, 33] have been used. In phase shifting method, field of two sheared and orthogonally polarized beams can be presented as

$$t_1 = a_1 e^{i(\theta_1 - \phi)} \quad (2.9)$$

$$t_2 = a_2 e^{i(\theta_2 + \phi)} \quad (2.10)$$

Where a is the amplitude and θ is the phase of the light pass through the specimen, 2ϕ is the prism-induced phase bias between two beams. $\Delta\theta = \theta_1 - \theta_2$ is the phase difference caused by the specimen phase gradient. The intensity of final image is

$$I = |t_1 + t_2|^2 = a_1^2 + a_2^2 + 2a_1a_2 \cos(\Delta\theta + 2\phi) \quad (2.11)$$

In DIC, phase bias can be changed by phase shifting. By incrementing 2ϕ by $\pi/2$ step by step, four images can be obtained. Then the specimen phase difference (phase gradient) can be obtained by

$$\Delta\theta = \tan^{-1} \frac{I_{\pi/2} - I_{3\pi/2}}{I_0 - I_\pi} \quad (2.12)$$

The next step is to recover phase from its gradient. Many methods have been adopted such as integration, filtering in Fourier domain, 6-frame, 4-frame, 2-frame algorithm, non-iterative and iterative computation.

Applications include refractive index analysis of optical fiber [33], cell imaging [34]. Figure 2.19 shows an example of cell imaging of quantitative DIC microscope. (a) is a DIC image of a cheek cell. (b) is phase reconstruction at a single plane from (a). (c) is 3D topological view of (b).

Quantitative DIC microscope is based on DIC microscope which is basically a shearing interferometer for phase imaging. Its geometry can be reflection and transmission. What it records is interference intensity image of DIC microscope. Reconstruction method from this image is first to extract phase gradient and then calculate phase from its gradient. Therefore it is an indirect phase measurement technique.

The advantages of this method: It utilizes conventional DIC microscope image as input for calculation. Thus the setup is simple.

The disadvantages of this method: complex calculation from phase gradient to phase, the specimen of DIC should be weak phase object which means the phase shift induced by specimen should be smaller than $\pi/2$. It is not easy to perform real time measurement, as more than one image needs to be recorded.

Figure 2.19 (a) DIC image of a cheek cell; (b) Phase reconstruction at a single plane from (a); (c) 3D topological view [35].

2.2.7 Spectral-Domain Phase Microscope

Spectral-domain phase microscope [32, 33, 35] is a phase-sensitive functional derivative of spectral-domain optical coherence tomography (OCT) to produce depth-resolved intensity and phase profiles with significantly improved phase stability compared to systems based on time domain OCT [36]. This technique is based on interference. It is also a quantitative phase measurement method. It can generate 3-D

quantitative phase-contrast image of a specimen simply by scanning the beam laterally as it measures phase profile in depth [37, 38].

Figure 2.20 shows a setup of spectral domain phase microscope which includes common path spectral-domain OCT. A broadband 840nm superluminescent diode (50nm FWHM) is used as light source. Swept laser source can also be used as light source in such technique [37].

Figure 2.20 (a) Schematics of spectral-domain phase microscopy. (b) Sample placed between a coverslip and a microscope slide [36].

As shown in Fig. 2.20 (b), the reflection from the top surface of a coverslip serves as the reference optical field, and the backscattered waves from the sample are the measurement fields. When the thickness of the coverslip is larger than that of the specimen—cell in this case, the interference signal referenced to the top surface of the coverslip can be distinguished and separated easily from interference signals referenced to other surfaces. Details of this setup are introduced in reference [37].

The information value directly recorded in this technique is interference intensity at point (x, y) where the beam locates on the specimen. It is in the form as Eq. (2.13) [37].

$$I(k) |_{(x,y)} = 2[R_r R_s(z)]^{1/2} S(k) \cos(2k\Delta p) |_{(x,y)} \quad (2.13)$$

where k is the free-space wave number; z is the geometrical distance; R_r and R_s are the reference reflectivity and specimen reflectivity at depth z , respectively; $S(k)$ is the source power spectral domain and Δp is the OPD between the reference and the sample beams. The spectral domain phase microscope is of reflection geometry.

To reconstruct phase from the recorded interference intensity, depth profile $F(z)$ of complex-valued is acquired by performing discrete Fourier transform with respect to $2k$ [37]. As interference component referenced to the top surface of the coverslip can be separated from interference components referenced to other surfaces, phase information which is a function of z can be obtained from the argument of $F(z)$:

$$\phi(z) |_{(x,y)} = \tan^{-1} \left\{ \frac{\text{Im}[F(z)]}{\text{Re}[F(z)]} \right\} = 2 \frac{2\pi}{\lambda_0} \Delta p(z) \quad (2.14)$$

where λ_0 is the center wavelength of the source. The depth-resolved phase measurement is performed as the beam scans laterally across the specimen point by point. We see the OPD at certain layer of depth z with respect to a reference lay—a coverslip this time can be obtained. Therefore the final information got from this technique could be quantitative phase or quantitative OPD which is optical thickness related to n and d where n is refractive index and d is the physical thickness of specimen.

Figure 2.21 Images of human epithelial cheek cells. (a) Image recorded by a Nomarski microscope (10X; N.A., 0.3); the bar represents 20 μm . (b) Spectral –domain optical coherence phase microscope image, along with the gray scale denoting the OPD in nanometres. (c) Surface plot of (b), showing optically thick structures such as nuclei and subcellular structures in the cell [38, 39].

Such technique has been applied in human epithelial cheek cells imaging [37], in vivo human retinal imaging [37], and cellular dynamics measurement [39], etc.

One example of applications is imaging of human epithelial cheek cells shown in Fig 2.21. (a) is image recorded by a Nomarski microscope, the bar represents 20 μm . (b) is spectral –domain optical coherence phase microscope image, along with the gray scale

denoting the OPD in nanometres. (c) is surface plot of (b), showing optically thick structures such as nuclei and sub-cellular structures in the cell.

No specimen preparation is required in such technique. Furthermore, phase information in certain depth inside the specimen can be isolated out from the whole specimen which means phase sectioning can be performed. However, this technique relies on single point measurement, which for imaging purposes required raster scanning. Thus it is time consuming.

2.2.8 Digital Holography

Digital holography is an interference based technique for quantitative phase imaging. It can simultaneously provide quantitative amplitude and phase images. It also has the capability to numerically reconstruct different object planes without using any optomechanical movement [40].

Digital holography was developed from conventional holography. In conventional holography, the interference between the coherent object and reference waves produces an interference pattern--hologram, which contains the information about not only the intensity of light but also its phase. Conventional holography uses a photographic plate to record the hologram and hologram is developed by photochemical processes. The hologram is then illuminated by the original reference wave. The original object optical field is reproduced by the propagation of diffracted light from the hologram. As the reproduced holographic image retains the information of not only the amplitude but also the phase of object wave, this image is the exact 3D replica of the original object.

As the conventional processes of holographic recording and hologram photochemical development are complicated and time-consuming, more interests and efforts have been shifted towards digital holography [36]. Main advantages of digital holography over conventional holography are listed below:

- Rapid image acquisition
- Accessibility to quantitative amplitude and phase information
- Various image processing techniques can be applied to the complex field.

Comparing with conventional holography, in digital holography, hologram is sampled by a high resolution CCD array and transferred into a computer as an array of numbers [41]. The recorded digital hologram is multiplied by the digital reference wavefield in the hologram plane and the digital diffracted field in the image plane, which is another numerical array of complex numbers, is determined by the Fresnel-Kirchoff integral to numerically calculate the intensity and the phase distribution of the reconstructed real image array [42-44]. Numerical reconstruction could be performed by Fresnel transform, Huygens convolution, and angular spectrum methods [45].

The principle of digital holography for 3D imaging includes two parts: digital recording and numerical reconstruction which is explained in Fig. 2.22. During the recording process, Fig. 2.22 (a), the illumination wave is scattered and reflected by the object and becomes the object wave at the object plane: $O_1(x, y) = o_1(x, y)\exp\{i\varphi_{o1}(x, y)\}$ where $o_1(x, y)$ is the amplitude and $\varphi_{o1}(x, y)$ is the phase. In this context, the 3D object profile information is encoded in the phase of object wave $\varphi_{o1}(x, y)$. This is because

that the object surface topography changes the optical path length of the illumination wave. The object wave propagates from the object plane to the recording device and becomes $O(\xi, \eta) = o(\xi, \eta)\exp\{i\varphi_o(\xi, \eta)\}$ at the recording plane where $o(\xi, \eta)$ is the amplitude and $\varphi_o(\xi, \eta)$ is the phase. Since the recording devices can only record intensity information, a reference wave $R(\xi, \eta) = r(\xi, \eta)\exp\{i\varphi_r(\xi, \eta)\}$ is used to encode the phase information $\varphi_o(\xi, \eta)$ into the interference pattern—hologram. The hologram $I(\xi, \eta)$ is expressed as

$$\begin{aligned} I(\xi, \eta) &= [R(\xi, \eta) + O(\xi, \eta)] \times [R(\xi, \eta) + O(\xi, \eta)]^* \\ &= o(\xi, \eta)^2 + r(\xi, \eta)^2 + o(\xi, \eta)r(\xi, \eta)\exp\{i[\varphi_o(\xi, \eta) - \varphi_r(\xi, \eta)]\} + \\ &\quad o(\xi, \eta)r(\xi, \eta)\exp\{-i[\varphi_o(\xi, \eta) - \varphi_r(\xi, \eta)]\} \end{aligned} \quad (2.15)$$

where * denote the complex conjugate. It can be seen that the phase $\varphi_o(\xi, \eta)$ is included in the third term of Eq. (2.15). The phase $\varphi_o(\xi, \eta)$ is thus encoded in the recorded intensity of the hologram.

The numerical reconstruction process is shown in Fig. 2.22 (b). We illustrate it with a reconstruction example in Fig. 2.23. The third term of Eq. (2.15) is extracted from the hologram in the spectrum domain as shown in Fig. 2.23 (a-c) and is combined with the conjugate of the reference wave R^* in space to obtain the term $o(\xi, \eta)r(\xi, \eta)^2\exp\{i\varphi_o(\xi, \eta)\}$. Usually the amplitude of the reference wave $r(\xi, \eta)$ is a constant value which can be ignored. Therefore the object wave at the hologram plane $O(\xi, \eta) = o(\xi, \eta)\exp\{i\varphi_o(\xi, \eta)\}$ can be obtained as in Fig. 2.23 (d). The object wave at the object plane $O_1(x, y) = o_1(x, y)\exp\{i\varphi_{o1}(x, y)\}$ of Fig. 2.23 (e) is achieved by back wave propagation from hologram plane to the object plane which is performed by a digital Fresnel transform. The amplitude image $o_1(x, y)$ and phase image $\varphi_{o1}(x, y)$

shown in Fig. 2.23 (f) and (g) are obtained. From the phase image $\varphi_{o1}(x, y)$, the OPD value is calculated. According to the relationship between OPD and the object profile in different modes, the reconstruction of 3D profile of Fig. 2.23 (h) is accomplished.

Fig. 2.22 The principle of digital holography for 3D imaging: (a) digital recording and (b) numerical reconstruction.

Figure 2.23 An example of the implementation of numerical reconstruction process.

Holographic configurations can be divided into off-axis geometry and in-line geometry [46, 47]. Our group has done leading research in the field of in-line digital holography [48-64]. The problem associated with in-line geometry is the overlapping of zero-order and twin images. Though methods to solve the problem exist, in-line digital holography is still not suitable for real time phase imaging due to the overlapping of different terms

and the sensitivity of phase. Therefore in the study of phase measurement, only off-axis configuration will be focused on.

Fig. 2.24 (a) is the schematic geometry of transmission digital holography. One coherent laser beam is split into two parts – the reference beam passes through the beam splitter and illuminates the CCD directly; the object beam illuminates the sample on the stage and passes through it. These two beams interfere at the CCD plane to generate the hologram. Lenses 1 and 2 are used for collimated light. The insets of Fig. 2.24 (a) and (b) show the off-axis geometry in detail.

The schematic geometry of reflection digital holography is shown in Fig. 2.24 (b). Laser beam passes through beam splitter and is divided into two parts. One is object beam which illuminates the specimen. The light reflected from the specimen then travels towards the CCD. The other is reference wave which is reflected by a mirror and then arrives at the CCD to interfere with object wave.

In digital recording, the sampling theory should be satisfied to fully resolve the interference pattern to further acquire reconstructed image of good quality. The hologram is then recorded and digitalized by CCD and transported into computer and saved as a digital hologram.

Figure 2.24 (a) Schematic geometry of transmission digital holography; (b) Schematic geometry of reflection digital holography.

In numerical reconstruction, two algorithms of wave propagation approaches, Fresnel and convolution, can be used in the propagation calculation based on diffraction theory. The way to numerically calculates this propagation by Fresnel transformation is shown below [48, 51-54, 56, 57, 59]:

$$O_1(m\Delta x, n\Delta y) = A \times \exp \left[j \frac{\pi}{\lambda d} (m^2 \Delta x^2 + n^2 \Delta y^2) \right] \times FFT \left\{ O(k, l) \exp \left[j \frac{\pi}{\lambda d} (k^2 \Delta \xi^2 + l^2 \Delta \mu^2) \right] \right\}_{m,n} \quad (2.16)$$

where $A = \frac{1}{j\lambda d} \exp \left(j \frac{2\pi}{\lambda} d \right)$; k, l, m, n are integers $(-N/2 \leq k, l, m, n \leq N/2)$; d is the reconstruction distance from the hologram plane to the image plane. $\Delta \xi$ and $\Delta \mu$ are sampling intervals in hologram plane and Δx and Δy are sampling intervals in image plane which are equal to

$$\Delta x = \Delta y = \frac{\lambda d}{N \Delta x} = \frac{\lambda d}{L} \quad (2.17)$$

where L is the size of CCD. If digital hologram is not padded in the numerical reconstruction, L is also the size of digital hologram. Δx and Δy defines the transverse resolution in the image plane. $O_1(m\Delta x, n\Delta y)$ is the reconstructed object wavefront at the image plane. It is an array of complex numbers whose modulus gives the amplitude and the arctangent of the imaginary over real part provides the phase. Phase unwrapping provides the absolute distribution of the phase.

The phase provides the OPD. For reflection geometry, OPD is proportional to surface profile of specimen. For transmission geometry, OPD is the integration of the product of the thickness of the specimen and the refractive index difference between the object medium and the host medium along the light propagation direction in specimen. Hence digital holography provides the surface profile apart from a constant related to refractive index provided the refractive index is constant across the entire specimen and in air as in Fig. 1.1 (b).

Figure 2.25 Images of a living mouse cortical neuron in culture: (a), dark PhC image; (b), DIC image; (c), raw image; (d), perspective image in false colors of the phase distribution obtained with the DHM [65-67].

Digital holography can also be integrated with microscope objective (MO) to increase lateral resolution for microscopic specimen image and measurement. This integration is also called digital holographic microscopy (DHM) [68-73].

Digital holography has been applied to surface profilometry [51, 53-57, 66, 67, 74-78] and quantitative visualization and measurement [40, 79-86]. An example of phase measurement application for cell imaging by transmission setup is shown in Fig. 2.25. Images of the same living neurons in culture obtained with a DHM and phase contrast microscope (PhC), and DIC microscopy are presented. (a) is dark PhC image. (b) is DIC

image. (c) is raw quantitative phase image obtained by the DHM image. By assuming a constant and homogeneous cellular refractive index (actually cell is not of homogeneous cellular refractive index), 3D perspective of living neurons is obtained as shown in Fig. 2.25 (d). Two scale bars presents the quantitative phase and thickness of the cell.

Fig. 2.26 Phase measurement results by reflection setup (a) 3D perspective of the height of a 1951 United States Air Force (USAF) target with aluminum deposition [67]; (b) 3D topography of steps with different heights [78].

An example of phase measurement application of a reflection setup is shown in Fig. 2.26. The 3D topography images of specimen height are presented in Fig. 2.26(b).

Digital holography is an optical technique for both intensity and phase imaging based on interferometry. But there are differences between digital holography and other phase imaging approaches based on interferometry introduced in this chapter, such as FPM, HPM and DPM. Those approaches record interferograms of image field and reference field and retrieve phase of image field from interferograms. However, CCD is placed before the image plane of object in digital holography so that it actually records

interferogram of the object wavefront away from the image plane and the reference wave. And in reconstruction process, object wavefront at the CCD plane is extracted from interferogram in the first step. Then the object wavefront at the CCD plane is propagated to the image plane. Thus, the complex wavefront of image is the final result of digital holography and from it the quantitative amplitude and phase image can easily be obtained.

As digital holography can reconstruct the wave propagation from CCD plane to image plane, not only the image plane but the whole complex optical field in the volume between CCD plane and image plane is calculated. Therefore DHM is capable to provide more information than other interferometry methods. And both reflection setup and transmission setup are available for different applications. Setups are relatively simple.

2.3 Conclusion

Digital holography shows better capabilities among all the phase imaging approaches. Comparing with phase microscope, DIC microscope and confocal fluorescence microscope, digital holography is quantitative phase measurement approach rather than qualitative. Compared to the 3D profiling confocal microscope which works by scanning and is time consuming, digital holography is fast and real-time and can provide nanometer rather than micrometer axial resolution. In comparison with quantitative phase microscopy, digital holography only needs to record one image instead of 3 images and the reconstruction method is much simpler. In relation to the

Fourier phase microscope, Hilbert phase microscope and diffraction phase microscope, digital holography system contains fewer optical elements and thus is less complex and easier to adjust and phase aberration due to optics is less severe. Digital holography is capable to provide more information than other interferometry methods in that not only the image plane but the whole complex optical field in the volume between CCD plane and image plane is calculated. Digital holography provides direct phase image instead of phase gradient image and thus less calculation is required as compared to the quantitative DIC method. Digital holography records only one image instead of 2-3 image for quantitative DIC. Furthermore, it is not limited to phase objects with phase shift less than $\pi/2$. The spectral-domain phase microscope is a raster scan approach while digital holography is the full field approach of digital holography. Also, the digital holography can both be used for opaque objects in reflection mode and for transparent objects in transmission mode.

Therefore, digital holography is chosen for advanced study of enhanced quantitative phase measurement. Further details of the digital holography technique are introduced in the following chapter. Table 2.1 concludes and compares the characteristics of all the phase measurement methods discussed in this chapter.

CHAPTER 2 LITERATURE REVIEW OF PHASE MEASUREMENT METHODS

Methods	Quantitative /Qualitative	Reflection / Transmission	Interference	Recording Content	Reconstruction Method	Spatial / Frequency Domain	Result	Specimen Preparation
Phase Microscope	Qualitative	Both	Yes	Intensity of the focused image	No	Spatial	Entangled image of amplitude with phase	No
DIC Microscope	Qualitative	Both	Yes	Intensity of the focused image	No	Spatial	Entangled image of amplitude with phase gradient	No
Fluorescence Confocal Microscope	Qualitative	Transmission	No	Intensity of the focused image	No	Spatial	Intensity	fluorescence d yed
3D profiling Confocal Microscope	Quantitative	Both	No	Intensity of different layers	Combination of intensity of different layers	Spatial	Tomography	No
Quantitative Phase microscope	Quantitative	Transmission	No	Intensity of focused image and defocused images	Transport of Intensity Equation	Spatial	Optical Thickness	No
Fourier Phase	Quantitative	Transmission	Yes	Interference patterns of phase-	Phase-shift interferometry	Spatial	Optical Thickness	No

CHAPTER 2 LITERATURE REVIEW OF PHASE MEASUREMENT METHODS

Microscope				shifted image field with a planar reference field	algorithm			
Hilbert Phase Microscope	Quantitative	Transmission	Yes	Interferogram of image field with a planar reference field	Hilbert Transform	Spatial	Optical Thickness	No
Diffraction Phase Microscope	Quantitative	Transmission	Yes	Interferogram of image field with a planar reference field	Hilbert Transform	Spatial	Optical Thickness	No
Quantitative DIC Microscope	Quantitative	Both	Yes	2-4 DIC microscope images	Phase shift interferometry algorithm for phase gradient extraction and then calculation algorithm to extract phase from its gradient	Spatial	Height Profile (reflection) / Optical Thickness (transmission)	No
Spectral-Domain phase microscope	Quantitative	Reflection from layers in depth	Yes	Interferogram of backscattered waves from the sample in certain depth and reference waves reflected from the top a coverslip	Fourier Transform	Frequency	Optical Thickness	No

CHAPTER 2 LITERATURE REVIEW OF PHASE MEASUREMENT METHODS

Digital Holography Microscope	Quantitative	Both	Yes	Interferogram of object wavefront with a planar reference field	Fourier transform and wave propagation based on diffraction theory	Spatial	Height Profile (reflection) / Optical Thickness (transmission)	No
--------------------------------------	--------------	------	-----	---	--	---------	--	----

*n is the refractive index of specimen and d is the thickness of specimen in this table.

Table 2.1 Characteristics of all the phase object imaging approaches discussed in Chapter 2.

CHAPTER 3 DIGITAL HOLOGRAPHY

In this chapter, a general introduction of digital holography (DH) is given. DH technique has two parts: digital recording and numerical reconstruction.

In the digital recording process, hologram is recorded using different optical configurations onto a CCD or other digital recording devices. The optical configurations for DH recording are introduced in section 3.1.

During the recording stage the sampling theory should be satisfied to resolve the interference pattern (hologram) for quality of the reconstructed image. The necessary conditions to achieve this are discussed in section 3.2.

In the numerical reconstruction stage the object wavefield of the hologram plane or CCD plane is obtained at the first step. Next the wavefield is propagated from the hologram/CCD plane to the image plane based on diffraction theory of wave propagation. Two algorithms of wave propagation, the Fresnel and the convolution methods, are discussed in section 3.3.

After numerical reconstruction, the complex object wavefield in digital form can be obtained. Direct manipulations of the deduced amplitude and phase for the numerical correction of aberrations are reviewed in section 3.4.

DH is capable of providing reconstructed images with a diffraction-limited lateral resolution down to a few hundreds of nanometers depending on the numerical aperture (NA) of the system and axial/phase resolution of 0.0349-0.0698 radians due to the interferometric nature of this method[40]. Section 3.5 gives a general review of previous studies on the lateral resolution, its improvement and axial resolution and accuracy.

Based on the review of current DH development, possible exploration directions are pointed out in section 3.6.

3.1 Configurations of DHM

There are different set-ups for recording holograms [65]. Herein, the set-ups are classified into two main generic types: in-line holography and off-axis. In-line holography has two modes. In the first mode only one illumination beam is used. Part of the beam scattered by the object serves as object beam and the other part of light unaffected by the object serves as reference beam. In the other mode, one beam is split into reference beam and object beam which illuminates the object. At the recording device, these two beams interfere with each other. There is no tilt angle between the two beams. In off-axis holography, one beam is split into reference beam and object beam. But there is a tilt angle between the reference beam and the object beam. The drawback of in-line holography is that both the undiffracted reference beam and object beam are in the same direction. The in-focus image is always overlapped by the out of focus image of the other one. Therefore the phase shifting technique is needed to determine the phase. The off-axis arrangement can overcome this drawback in nature.

CHAPTER 3 DIGITAL HOLOGRAPHY

Holography can also be classified into reflection mode and transmission mode according to the different ways that light interacts with the object. If the optical information recorded by the hologram is reflected by the object, the system is characterized as reflection holography. If the optical information recorded by the hologram is transmitted through the object, the system is characterized as transmission holography.

Based on the distance between the object and hologram or between the focused object image and hologram when imaging lens is used, holography can be divided into image holography, Fresnel holography and Fraunhofer holography. If the distance between the imaged object and the hologram is zero, it is image holography system. If the distance is in the Fresnel diffraction region, it is Fresnel holography system. If the distance is in the Fraunhofer diffraction region, it is Fraunhofer holography system.

Holography can be divided into lensless holography and holography with lens. If no lens is used in the system or lenses are used but not to manipulate the object wave after it interacts with object, the system is a lensless holography system. If lens or microscope objective is used to generate object image and to improve the lateral resolution, the system can be called digital holography microscope. If the reference beam and object illumination beam in lensless holography are both divergent with the same curvature, this is a generalized lensless holographic microscope which can also provide lensless magnification. The reconstructed image is laterally magnified by $\frac{z_r}{z_r - z_s}$ times where z_r is

CHAPTER 3 DIGITAL HOLOGRAPHY

the distance between light source plane to hologram plane and z_s is the distance from object plane to hologram plane.

This thesis deals with digital Fresnel off-axis lensless holography using beams of identical curvatures. The schematic set-ups are shown in Fig. 3.1 where (a) is transmission mode and (b) is reflection mode.

In transmission set-up shown in Fig. 3.1 (a), a coherent laser beam is split into two parts – the reference beam passes through the beam splitter and illuminates the CCD directly; the object beam illuminates the sample on the stage. Lens 1 is adjusted to generate object wave with desired curvature for illuminating the specimen. After the object beam passes through the specimen, the two beams interfere at the CCD plane to generate the hologram. Lens 2 is used to adjust the curvature of reference beam to produce straight interference pattern with object beam at the CCD.

In the reflection setup shown in Fig. 3.1 (b), the position of lens is adjusted to generate wave with the desired curvature. The wave is split into reference wave and object beam by beam splitter. The object beam illuminates the specimen. The light reflected from the specimen travels towards the CCD. The reference beam is reflected by a mirror and arrives at the CCD to interfere with the object wave. The position and tilt angle of the mirror can be adjusted to generate high contrast interference pattern with the desired fringe direction and frequency.

Figure 3.1 schematic of DH in (a) transmission mode and (b) in reflection mode.

At the CCD, the interference between the object wave **O** and the reference wave **R** creates the hologram:

$$I(\xi, \eta) = |R|^2 + |O|^2 + R^*O + RO^* \quad (3.1)$$

where * denote the complex conjugate. R is the reference wave with $R = |r|exp\{i\varphi_r\}$ where r is the amplitude and φ_r is the phase of the reference wave. O is the object wave

with $O = |o| \exp\{i\varphi_o\}$ where o is the amplitude and φ_o is the phase of the reference wave.

The hologram is sampled by the CCD array and then transferred into a computer as an array of numbers. This digital hologram is multiplied by a digital reference wave to reconstruct the object wavefield at the hologram plane. The diffracted field at the image plane is then determined using Fresnel-Kirchoff diffraction integral[45] to obtain the numerical intensity and phase distribution at the image plane.

In the following, digital recording and numerical reconstruction approaches of hologram are discussed in greater detail.

3.2 Digital Recording

The CCD in the DH system as in Fig. 3.1 records the hologram which is given by Eq. (3.1). If we assume a planar reference wave $R(\zeta, \eta) = \sqrt{I_R} \exp\{ik \sin \theta \zeta\}$ where I_R is its intensity, the hologram is expressed as below [87] :

$$I(\zeta, \eta) = I_R + |O(\zeta, \eta)|^2 + \sqrt{I_R} \exp\{-ik \sin \theta \zeta\} O(\zeta, \eta) + \sqrt{I_R} \exp\{ik \sin \theta \zeta\} O(\zeta, \eta)^* \quad (3.2)$$

In the spatial domain, the phase factor $\exp\{-ik \sin \theta \zeta\}$ in the third term indicates that the virtual image is deflected by an angle $-\theta$. The phase factor $\exp\{ik \sin \theta \zeta\}$ in the fourth term deflects the real image by an angle θ . But the phase factor $\exp\{ik \sin \theta \zeta\}$ or $\exp\{-ik \sin \theta \zeta\}$ does not affect the zero order terms (the first two terms in Eq. (3.2)).

The geometry of the DH recording is shown in Fig. 3.2 with reference wave $R(\zeta, \eta) = \sqrt{I_R} \exp\{ik \sin \theta \zeta\}$. If the CCD has $N \times M$ pixels with distances $\Delta \xi$ and $\Delta \eta$ between pixel centers in the x -direction and y -direction, respectively, with the law of generality, we examine the array of pixels in one direction. Thus CCD size is $2D = N\Delta \xi$. The object of size L is symmetrically located along the x -axis and perpendicular to the optical axis z -axis. The plane reference wave $R(\zeta, \eta) = \sqrt{I_R} \exp\{ik \sin \theta \zeta\}$ impinges onto the CCD with a tilt angle θ . We first focus on the object wave spectrum collected by the CCD array. The spectrum of object wave collected by the CCD with size $2D$ from an arbitrary point $x = x_i$ of object is shown in Fig. 3.3 (a). The spectrum extends $(\frac{x_i}{\lambda z} - \frac{D}{\lambda z}, \frac{x_i}{\lambda z} + \frac{D}{\lambda z})$ according to the finite chirp function properties [45]. Its corresponding bandwidth at the hologram plane is $\frac{2D}{\lambda z}$ as shown in Fig. 3.3 (a). As the spectrum for the entire object collected by the CCD is the sum of the spectrum of all the points along its extent collected by the CCD according to the linearity theorem of Fourier transform, it extends from $-\frac{D}{\lambda z} - \frac{L}{2\lambda z}$ to $\frac{D}{\lambda z} + \frac{L}{2\lambda z}$ as depicted in Fig. 3.3 (b). The resulting bandwidth is $\frac{2D+L}{\lambda z}$. The interference term $R(\zeta, \eta)^* O(\zeta, \eta) = \sqrt{I_R} \exp\{-ik \sin \theta \zeta\} O(\zeta, \eta)$ in Eq. (3.2) has a spectrum as shown in Fig. 3.3 (c) which is shifted by the carrier frequency f_c with $f_c = \frac{\sin \theta}{\lambda}$ from Fig. 3.3 (b). Using the paraxial approximation [45], the carrier frequency f_c can be expressed by $f_c = \frac{\theta}{\lambda}$ which is related to tilt angle θ .

According to the sampling theory, the largest bandwidth which can be recorded by CCD is determined by $2f_{max} = \frac{1}{\Delta \xi}$ as in Fig. 3.3 (c). Therefore $\Delta \xi$ and $\Delta \eta$ play a crucial role in determining the system resolution. If the object bandwidth exceeds this limitation, the

spectra would overlap and high frequency information of the object is lost. In order to avoid aliasing, the following condition should be satisfied:

$$\frac{2D+L}{\lambda d} \leq \frac{1}{\Delta\xi} \quad (3.3)$$

For a CCD array with fixed pixel size and extent, Eq. (3.3) actually sets upper limits for the object extent L and sets a lower limit for the distance d between object and CCD array.

Figure 3.2 Geometry for recording off axis digital Fresnel hologram[45].

Figure 3.3 Spectra of points along object with extent L at the CCD plane. (a) The spectrum of an arbitrary point x of the object arriving at the CCD; (b) The spectrum of the whole object with extent L arriving at the CCD plane; (c) The spectrum of the whole object with extent L shifted by the carrier frequency due to the tilt angle of off-axis geometry in the hologram.

Besides sampling effect, the requirement on the reference wave angle θ to separate the zero order and the two first orders in Eq. (3.2) has been discussed in references [48, 53, 88, 89]. Those ideas can be expressed with Fig. 3.4 where the spectrum of zero order and the two first orders in Eq. (3.2) are shown. In order to fully separate them as seen in Fig. 3.4, $-\frac{D+L/2}{\lambda d} + f_c$ should be larger than $\frac{2D+L}{\lambda d}$. This requirement sets a lower limit for the tilt angle θ between the reference wave and object wave. The carrier frequency f_c should be

$$f_c \geq \frac{3 \times 2D + 3L}{2\lambda d}. \quad (3.4)$$

And the angle θ between reference and object beams should be

$$\theta \geq \frac{3 \times 2D + 3L}{2d}. \quad (3.5)$$

Figure 3.4 Spectrum distribution of a hologram [89].

But the requirements in Eq. (3.4) and (3.5) are not necessary in certain cases as shown in Fig. 3.5. As in Fig. 3.5, in the x -direction, the carrier frequency f_c does not satisfy Eq. (3.4). But as spectrum is of two dimensions, the zero order does not overlap with the first order in the x -direction. Hence, in some cases, Eq. (3.4) and (3.5) are not needed to be satisfied. But Eq. (3.3) should be followed so that the sampling theorem is fulfilled. Here we provide an example of a typical DH system with two planar waves. The parameters of the system are shown below:

- CCD size: $2D = 5.952mm$
- Wavelength: $\lambda = 633nm$
- Pixel size: $\Delta\xi = 4.65\mu m$

According to the above system parameters, the best resolution of system in Eq. (3.3) is $4.65\mu m$. For a point object with infinitely small extent $L \rightarrow 0$, the recording distance d should satisfy the condition $d \geq 43.72mm$ under the requirement of Eq. (3.3). For a

fixed recording distance $d = 100\text{mm}$, according to Eq. (3.3), the object size L should follow the condition $L \leq 7.66\text{mm}$.

Figure 3.5 An example of hologram spectrum with zero order and two first orders.

3.3 Numerical Reconstruction

Numerical reconstruction simulates the process of optical reconstruction. Two procedures are involved in the numerical reconstruction--obtaining the object wave field at the hologram plane from the hologram and object wavefront propagation from hologram plane to image plane.

3.3.1 Object Wavefield Reconstruction at Hologram Plane

In this section we adopt the coordinate system of digital holography as shown in Fig. 3.6.

Figure 3.6 Coordinate system for digital holography[41].

In the process of recording and reconstruction of digital holography, hologram could be written as Eq. (3.1) and Eq. (3.2) using the coordinate systems shown in Fig. 3.6.

As discussed in the last section, the last two first order diffraction terms in Eq. (3.2) propagate along different directions and hence they can be observed separately. If we reconstruct the hologram by illuminating it with a planar wave perpendicular to the hologram plane which corresponds to multiplying the hologram with 1 in numerical reconstruction, the term $\sqrt{I_R} \exp\{-ik \sin \theta \zeta\} O(\zeta, \eta)$ propagates in the direction of $-\theta$ and the term $\sqrt{I_R} \exp\{ik \sin \theta \zeta\} O(\zeta, \eta)^*$ propagates in the direction of θ . The problem of such reconstruction is that the reconstructed image is not at the center of reconstruction plane and its distance from the center of the image plane differs with different reconstruction distances d . Therefore it is not convenient to find the image.

In practical numerical reconstruction, the hologram is multiplied with a numerical reference wave or its conjugate to acquire the object wavefront or its conjugate at the hologram plane. Here we give an example when the reference wave is $R(\zeta, \eta) =$

$\sqrt{I_R} \exp\{ik \sin \theta \zeta\}$ to illustrate this action. The product of hologram and the numerical reference wave becomes

$$\begin{aligned} R(\zeta, \eta) I(\zeta, \eta) &= I_R^{\frac{3}{2}} \exp\{ik \sin \theta \zeta\} + |O(\zeta, \eta)|^2 \sqrt{I_R} \exp\{ik \sin \theta \zeta\} + I_R O(\zeta, \eta) \\ &\quad + I_R [\exp\{ik \sin \theta \zeta\}]^2 O(\zeta, \eta)^* \end{aligned} \quad (3.6)$$

The product of the hologram and the conjugate of the numerical reference wave becomes

$$\begin{aligned} R(\zeta, \eta)^* I(\zeta, \eta) &= \\ I_R^{\frac{3}{2}} \exp\{-ik \sin \theta \zeta\} &+ |O(\zeta, \eta)|^2 \sqrt{I_R} \exp\{-ik \sin \theta \zeta\} + I_R [\exp(-ik \sin \theta \zeta)]^2 O(\zeta, \eta) + \\ I_R O(\zeta, \eta)^* &\end{aligned} \quad (3.7)$$

With Eq. (3.6), the modulation of the conjugate reference wave on the wavefront $O(\zeta, \eta)$ is eliminated in the third term. With Eq. (3.7), the modulation of reference wave on the conjugate of object wavefront $O(\zeta, \eta)^*$ is eliminated in the fourth term. Therefore, by wave propagation, a virtual image located at the center of image plane which corresponds to the position initially occupied by the object can be acquired from the third term of Eq. (3.6). And a real image located at the center of image plane on the opposite side of the hologram plane can be acquired from the fourth term of Eq. (3.7).

In the case when the curvatures of reference and object waves are identical, the tilt angle between the two waves can be compensated numerically without a prior knowledge of the tilt angle. This demodulation of the tilt due to off-axis geometry is done by moving the filtered first order component to the centre of the Fourier plane. We provide an example of this demodulation of the tilt at hologram plane as in Fig. 3.7. The spectrum

of the hologram is shown in Fig. 3.7 (a), the zero order term is located in the centre of the Fourier plane. The tilt angle θ leads to a translation of the first order spectrum by $-\frac{\sin\theta}{\lambda}$ and $\frac{\sin\theta}{\lambda}$ which makes the two first orders locate symmetrically with respect to the centre of the Fourier plane. In the demodulation of tilt, one of the first orders is digitally extracted from the hologram spectrum in Fig. 3.7 (b). Filtering can be performed digitally by multiplication of Fourier transform of the hologram by a defined digital mask with a transparent window centered at the carrier frequency $\frac{\sin\theta}{\lambda}$ of the first order. The center of the first order can be found at the point with highest intensity within the range of the first order spectrum in Fig. 3.7 (a). Thus the zero order and the twin image are eliminated using this digital mask. To eliminate the tilt, the filtered spectrum is shifted to the centre of the spectral space. The inverse Fourier transform of this translated spectrum provides the object wavefield without the tilt at the hologram plane. Hence the tilt angle between the reference and the object waves is compensated. Finally, after wavefront propagation from hologram plane to image plane, the image is located at centre of image plane as shown in Fig. 3.7 (c). Without this compensation of tilt, the image is not located at the center of image plane as shown in Fig. 3.7 (d). Other advantages are the elimination of zero order and twin images and enhancement of signal-noise ratio[87].

However in the case the curvatures of the reference and object waves are not identical, besides the tilt, the curvature difference of the reference wave from the object wave needs to be eliminated from the third or the fourth term of Eq. (3.1) to get correct reconstructed image. In such a case, prior knowledge of the curvature difference is needed. Otherwise other approaches for curvature and aberration compensation are

needed which will be discussed in section 3.4. Wavefront propagation from the hologram plane to image plane will be discussed in the next section.

Figure 3.7 Elimination of tilt induced by off-axis geometry. (a) Fourier spectrum of the original hologram; (b) Filtered Fourier spectrum of one first order; (c) numerically reconstructed amplitude image with tilt compensated; (d) numerically reconstructed amplitude image without tilt compensation.

3.3.2 Object Wavefront Propagation from Hologram Plane to Image Plane

Having obtained the complex object wavefront O at the hologram plane herein wavefront propagation from hologram plane to image plane is discussed. This propagation can be described using diffraction theory[45]. With the coordinate system of Fig. 3.8, the diffraction of a light wave at the hologram plane is described by the Huygens-Fresnel principle or Fresnel-Kirchhoff formula [45] as

$$\Gamma(\xi, \eta) = \frac{1}{j\lambda} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} O(x, y) \frac{\exp(j\frac{2\pi}{\lambda}\rho)}{\rho} \cos\theta dx dy \quad (3.8)$$

With

$$\rho = \sqrt{(x - \xi)^2 + (y - \eta)^2 + d^2} \quad (3.9)$$

where $O(x, y)$ is the object wavefront at hologram plane and ρ is the distance between a point in the hologram plane and a point in the reconstructed image plane, as shown in Fig. 3.8. The angle θ is also shown in Fig. 3.8. From geometry,

$$\cos\theta = \frac{d}{\rho} \quad (3.10)$$

Substituting Eq. (3.10) into Eq. (3.8), $\Gamma(\xi, \eta)$ can be written as

$$\Gamma(\xi, \eta) = \frac{d}{j\lambda} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} O(x, y) \frac{\exp(j\frac{2\pi}{\lambda}\rho)}{\rho^2} dx dy \quad (3.11)$$

Figure 3.8 Rectangular coordinate system of light diffraction[45].

Eq. (3.8) is the basis for numerical hologram reconstruction. As the reconstructed wavefield $\Gamma(\xi, \eta)$ is a complex function, both the intensity as well as the phase can be determined quantitatively[45]. This is in contrast to the case of optical hologram reconstruction which provides qualitative information.

- **Fresnel approximation of Huygens-Fresnel principle**

The Fresnel approximation[42] is based on the binomial expansion of Eq. (3.9). When b is less than 1, the binomial expansion of $\sqrt{1+b}$ can be expressed by

$$\sqrt{1+b} = 1 + \frac{1}{2}b - \frac{1}{8}b^2 + \dots \quad (3.12)$$

Apply Eq. (3.12) to Eq. (3.9)

$$\begin{aligned} \rho &= d \sqrt{1 + \left(\frac{x-\xi}{d}\right)^2 + \left(\frac{y-\eta}{d}\right)^2} \\ &= d + \frac{(x-\xi)^2}{2d} + \frac{(y-\eta)^2}{2d} - \frac{1}{8} \frac{[(x-\xi)^2 + (y-\eta)^2]^2}{d^3} + \dots \end{aligned} \quad (3.13)$$

For ρ^2 in the denominator of Eq. (3.11), the error by retaining only the first term d is generally acceptable. But for ρ in the exponent of Eq. (3.11), this approximation needs more careful consideration since ρ is multiplied by $\frac{2\pi}{\lambda}$ which is a large value. If the values of x and y values as well as ξ and η are small compared to the distance d between the hologram plane and image plane such that the maximum phase change induced by ignoring the $-\frac{1}{8} \frac{[(x-\xi)^2 + (y-\eta)^2]^2}{d^3} + \dots$ is less than 1 radian, then the following approximation can be used

$$\rho \approx d + \frac{(x-\xi)^2}{2d} + \frac{(y-\eta)^2}{2d} \quad (3.14)$$

This is equivalent to

$$d^3 \square \frac{\pi}{4\lambda} [(x-\xi)^2 + (y-\eta)^2]_{\max}^2 \quad (3.15)$$

These two approximations reduces Eq. (3.11) to

$$\Gamma(\xi, \eta) = \frac{\exp(j\frac{2\pi}{\lambda}d)}{j\lambda d} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} O(x, y) \exp\left\{j\frac{\pi}{\lambda d}[(x - \xi)^2 + (y - \eta)^2]\right\} dx dy \quad (3.16)$$

Eq. (3.16) can be rewritten as

$$\Gamma(\xi, \eta) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} O(x, y) g_F(x - \xi, y - \eta) dx dy \quad (3.17)$$

Eq. (3.17) is a convolution between $O(x, y)$ and $g_F(x, y)$ where the convolution kernel is

$$g_F(x, y) = \frac{\exp(j\frac{2\pi}{\lambda}d)}{j\lambda d} \exp\left[j\frac{\pi}{\lambda d}(x^2 + y^2)\right] \quad (3.18)$$

Alternatively, Eq. (3.11) can also be written as

$$\Gamma(\xi, \eta) = \frac{\exp(j\frac{2\pi}{\lambda}d)}{j\lambda d} \exp\left[j\frac{\pi}{\lambda d}(\xi^2 + \eta^2)\right] \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} O(x, y) \exp\left[j\frac{\pi}{\lambda d}(x^2 + y^2)\right] \exp\left[-j\frac{2\pi}{\lambda d}(x\xi + y\eta)\right] dx dy \quad (3.19)$$

We refer to both Eq. (3.16) and (3.19) as the *Fresnel diffraction integral* [45]. This approximation is valid for distance d that satisfies Eq. (3.15) which is said to be in the region of Fresnel diffraction, or equivalently in the near field of aperture.

These Fresnel diffraction integral enables reconstruction of the wavefield in image plane. The intensity is

$$I(\xi, \eta) = |\Gamma(\xi, \eta)|^2 \quad (3.20)$$

The phase is calculated by

$$\varphi(\xi, \eta) = \arctan \frac{\text{Im}[\Gamma(\xi, \eta)]}{\text{Re}[\Gamma(\xi, \eta)]} \quad (3.21)$$

where Re denotes the real part and Im the imaginary part.

- **Reconstruction by Fresnel transformation**

Eq. (3.19) can be numerically evaluated using the Fourier transform as [44, 45]

$$\Gamma(\nu, \mu) = \frac{\exp(j\frac{2\pi}{\lambda}d)}{j\lambda d} \exp[j\pi\lambda d(\nu^2 + \mu^2)] \mathcal{F} \left\{ O(x, y) \exp \left[j \frac{\pi}{\lambda d} (x^2 + y^2) \right] \right\}_{\nu=\frac{\xi}{\lambda d}, \mu=\frac{\eta}{\lambda d}} \quad (3.22)$$

Discrete formulation of Eq. (3.22) can be derived as [41, 66, 88].

$$\Gamma(m, n) = A \times \exp \left[j \frac{\pi}{\lambda d} (m^2 \Delta \xi^2 + n^2 \Delta \mu^2) \right] \times FFT \left\{ O(k, l) \exp \left[j \frac{\pi}{\lambda d} (k^2 \Delta x^2 + l^2 \Delta y^2) \right] \right\}_{m,n} \quad (3.23)$$

where $= \frac{\exp(j\frac{2\pi}{\lambda}d)}{j\lambda d}$, k, l, m, n are integers ($-N/2 \leq k, l, m, n \leq N/2$) and $O(k, l)$ is the digital wavefront at hologram plane. Sampling intervals $\Delta \xi$ and $\Delta \mu$ in the observation plane can be deduced as [66].

$$\Delta \xi = \Delta \eta = \frac{\lambda d}{N \Delta x} = \frac{\lambda d}{L} \quad (3.24)$$

where L is the size of CCD and also hologram.

• Reconstruction by Convolution approach

The propagation of light from the x - y plane to a parallel η - ξ plane at non-zero distance d in Fig. 3.8 is described by the Fresnel-Kirchhoff formula:

$$\Gamma(\xi, \eta) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} O(x, y) g(x - \xi, y - \eta) dx dy \quad (3.25)$$

where the impulse response $g(x - \xi, y - \eta)$ is given by

$$g(x - \xi, y - \eta) = \frac{1}{j\lambda} \frac{\exp \left(j \frac{2\pi}{\lambda} \sqrt{d^2 + (x - \xi)^2 + (y - \eta)^2} \right)}{\sqrt{d^2 + (x - \xi)^2 + (y - \eta)^2}} \quad (3.26)$$

where $\cos \theta = \frac{d}{\rho} \approx 1$, since for the θ required by sampling theorem, the resulting $\cos \theta$

differs by less than 1/1000 from 1 [45].

Eq. (3.26) shows that the linear system characterized by $g(x - \xi, y - \eta)$ is space-invariant: thus the integral of Eq. (3.25) is a convolution:

$$\Gamma(\xi, \eta) = O(x, y) \otimes g(x, y) \quad (3.27)$$

where

$$g(x, y) = \frac{1}{j\lambda} \frac{\exp\left(j\frac{2\pi}{\lambda}\sqrt{d^2+x^2+y^2}\right)}{\sqrt{d^2+x^2+y^2}} \quad (3.28)$$

According to convolution theorem, we have

$$\Gamma(\xi, \eta) = \mathcal{F}^{-1}\{\mathcal{F}\{O(x, y)\} \times \mathcal{F}\{g(x, y)\}\} \quad (3.29)$$

$\mathcal{F}$ denotes Fourier transform and $\mathcal{F}^{-1}$ is inverse Fourier transform. Thus based on Eq. (3.29), three Fourier transforms needs to be calculated for the propagation.

From the angular spectrum view point [90], the transfer function G is

$$G(\nu, \mu) = \mathfrak{F}\{g(x, y)\} = \exp\left[i2\pi\frac{d}{\lambda}\sqrt{1-(\lambda\nu)^2-(\lambda\mu)^2}\right] \quad \nu^2 + \mu^2 < \frac{1}{\lambda^2} \quad (3.30)$$

This saves one Fourier transform for reconstruction calculation:

$$\Gamma(\xi, \eta) = \mathcal{F}^{-1}\{\mathcal{F}\{O(x, y)\} \times G(\nu, \mu)\} \quad (3.31)$$

If we take a look at Fresnel approximation of Eq. (3.17) and (3.18), we recognize that Eq. (3.17) also has the form of a convolution with the convolution kernel $g_F(x, y)$. The transfer function of it in Fresnel approximation is

$$G_F(\nu, \mu) = \exp\left(j\frac{2\pi}{\lambda}d\right)\exp[-j\pi\lambda d(\nu^2 + \mu^2)] \quad (3.32)$$

The pixel sizes of the images reconstructed by convolution approach are equal to that of the hologram:

$$\Delta\xi = \Delta x \text{ and } \Delta\eta = \Delta y \quad (3.33)$$

The pixel sizes are therefore different from those of Fresnel approximation Eq. (3.24).

Now we can use directly the transfer function $G(m, n)$ or its Fresnel approximation $G_F(m, n)$ to perform reconstruction by $\Gamma = \mathcal{F}^{-1}\{\mathcal{F}\{O\} \times G\}$ or $\Gamma = \mathcal{F}^{-1}\{\mathcal{F}\{O\} \times G_F\}$. The whole process requires totally two Fourier transforms, which are effectively carried out using the FFT algorithm.

In total, there are four ways to reconstruct a digital hologram by the convolution approach[45]. We may define the exact impulse response g or its Fresnel approximation g_F and calculate $\mathcal{F}\{g\}$ or $\mathcal{F}\{g_F\}$. In the same way we may use directly the exact transfer function $G(m, n)$ or its Fresnel approximation $G_F(m, n)$.

Reconstruction of hologram by convolution approach results in a reconstructed image with same size for all reconstruction distances. The resulting image covers the area $N\Delta x \times N\Delta y$ of the scene instead of $\frac{\lambda d}{\Delta x} \times \frac{\lambda d}{\Delta y}$ with Fresnel transform method of reconstruction. As long as $d > N\Delta x^2 / \lambda$, the area of the reconstructed image by Fresnel transform is larger than that by convolution approach [65].

• COMPARISONS OF THE RECONSTRUCTION METHODS

The different implementations of the diffraction integral are summarized in table 3.1. There is a conceptual difference between Fresnel transform using the chirp function in the first row of the table and the next four methods: If we take the plane of the digital hologram as the spatial domain, then the first procedure gives a result in the spatial frequency domain due to only one Fourier transform. The other four algorithms which

can be named as convolution methods consist of a multiplication of the spectrum of O with a transfer function in the spatial frequency domain and an inverse Fourier transform back into the spatial domain. A consequence of this difference is the dissimilarity of the pixel size in the reconstructed images as in Eq. (3.24) and Eq. (3.33).

The pixel size $\Delta\xi \times \Delta\eta = \frac{\lambda d}{N\Delta x} \times \frac{\lambda d}{N\Delta y}$ in the Fresnel case using the chirp function

depends on the wavelength λ and the reconstruction distance d , while in the other four cases the pixel size is independent of those parameters. This makes the latter four algorithms useful if the reconstructed images have to be evaluated in different depths. The sizes of all reconstructions agree with each other and a direct comparison is allowed.

If a small object is located in the region where the Fresnel approximation is not suitable, the convolution approaches of the last two algorithms in Table 3.1 are recommended since they yield an exact solution to the diffraction integral as far as the sampling theorem is not violated. If the whole possible field of view for opaque or transparent objects has to be reconstructed, the Fresnel transform which is the first algorithm in Table 3.1 is better choice. There are also differences in the numerical treatment of the last four algorithms in table 3.1 which are discussed in reference.

As a conclusion, digital recording and numerical reconstruction of holograms offer new possibilities to optical metrology. We see that in wavefront acquisition at hologram plane, digital processing of hologram makes filtering the zero order and twin image possible. And the tilt due to off-axis geometry which could not be eliminated in classic optical holography can be compensated in DH. To calculate wavefront propagation

CHAPTER 3 DIGITAL HOLOGRAPHY

from hologram plane to image plane, numerical Fresnel transform or convolution methods can be used to evaluate the diffraction of wavefront.

In conclusion, by digital recording and numerical reconstruction of hologram, we obtain the complex object wavefield at the image plane. However, the complex wavefield contains the phase deformation/aberrations and other imperfections of optics. In order to acquire the correct complex wavefield of object, compensation of these phase deformations are required and necessary. In next section, we will discuss the approaches for compensating phase aberrations.

CHAPTER 3 DIGITAL HOLOGRAPHY

Method	Algorithm	Impulse Response/Transfer function	Pixel size of Reconstructed image
Fresnel-Approx. (Chirp function)	$\mathfrak{Z}^{-1}\{O \cdot R \cdot c\}$	$z(m, n) = \exp[j \frac{\pi}{\lambda d} (m^2 \Delta \xi^2 + n^2 \Delta \eta^2)]$ $c = \exp[j \frac{\pi}{\lambda d} (k^2 \Delta x^2 + l^2 \Delta y^2)]$	$\Delta \xi = \frac{\lambda d}{N \Delta x} = \frac{\lambda d}{L}$ $\Delta \eta = \frac{\lambda d}{N \Delta y} = \frac{\lambda d}{L}$
Fresnel-Approx. (Impulse response)	$\mathfrak{Z}^{-1}\{\mathfrak{Z}\{O \cdot R\} \cdot \mathfrak{Z}\{g_F\}\}$	$g_F(k, l) = \frac{\exp(j \frac{2\pi}{\lambda} d)}{j \lambda d} \exp[j \frac{\pi}{\lambda d} ((k \Delta x)^2 + (l \Delta y)^2)]$	$\Delta \xi = \Delta x$ $\Delta \eta = \Delta y$
Fresnel-Approx. (Transfer function)	$\mathfrak{Z}^{-1}\{\mathfrak{Z}\{O \cdot R\} \cdot G_F\}$	$G_F(m, n) = \exp(j \frac{2\pi}{\lambda} d) \exp[-j \pi \lambda d (\frac{m^2}{N^2 \Delta x^2} + \frac{n^2}{N^2 \Delta y^2})]$	$\Delta \xi = \Delta x$ $\Delta \eta = \Delta y$
Diffraction integral (Impulse response)	$\mathfrak{Z}^{-1}\{\mathfrak{Z}\{O \cdot R\} \cdot \mathfrak{Z}\{g\}\}$	$g(k, l) = \frac{1}{j \lambda} \frac{\exp(j \frac{2\pi}{\lambda} \sqrt{d^2 + (k \Delta x)^2 + (l \Delta y)^2})}{\sqrt{d^2 + (k \Delta x)^2 + (l \Delta y)^2}}$	$\Delta \xi = \Delta x$ $\Delta \eta = \Delta y$
Diffraction integral (Transfer function)	$\mathfrak{Z}^{-1}\{\mathfrak{Z}\{O \cdot R\} \cdot G\}$	$G(m, n) = \exp[i 2 \pi \frac{d}{\lambda} \sqrt{1 - (\lambda \frac{m}{N \Delta x})^2 - (\lambda \frac{n}{N \Delta y})^2}]$	$\Delta \xi = \Delta x$ $\Delta \eta = \Delta y$

Table 3.1[90]

3.4 Aberration Compensation

A typical phase aberration is associated with the use of a microscope objective (MO).

Figure 3.9 Configuration of the object arm in holographic microscopy with MO [66].

As shown in Fig. 3.9, the optical arrangement of the object arm is actually a single-lens imaging system. The MO produces a magnified image of the object. The hologram plane (CCD plane) is located between the MO and the image plane at a distance d from the image. This situation is equivalent to a no MO configuration with an object wave emerging directly from the magnified image in the reconstruction of intensity image. In the reconstruction process, the image will be in focus when the reconstruction distance d is the distance between CCD and the image during hologram recording.

The problem of reconstruction in the system with MO comes in phase imaging. Though the application of MO increases transverse resolution, it also induces wavefront distortion as shown in Fig. 3.10. This distortion influences only the phase but not the amplitude image.

Figure 3.10 Schematic of wave-front distortion by MO[66].

This curvature distortion and also other types of distortions can be compensated in the numerical reconstruction process by multiplying the reconstructed complex wavefront with the computed complex conjugate of phase aberration also called a digital phase mask (DPM). The digital phase mask to compensate MO distortion can be expressed in the form[66]

$$\Phi(m, n) = \exp\left[\frac{-i\pi}{Dis}(m^2\Delta\xi^2 + n^2\Delta\eta^2)\right] \quad (3.34)$$

where m, n are integers, $\Delta\xi$ and $\Delta\eta$ are the sampling intervals in observation plane as discussed in previous sections in Fig. 3.8, and Dis is

$$\frac{1}{Dis} = \frac{1}{d_i} \left(1 + \frac{d_o}{d_i}\right) \quad (3.35)$$

where d_o and d_i are the distances between object and MO, and image and MO respectively. Then the complete digital expression of Fresnel reconstruction algorithm of Eq. (3.23) becomes

$$\Gamma(m, n) = A \times \Phi(m, n) \exp\left[j \frac{\pi}{\lambda d} (m^2\Delta\xi^2 + n^2\Delta\eta^2)\right] \times FFT\left\{O(k, l) \exp\left[j \frac{\pi}{\lambda d} (k^2\Delta x^2 + l^2\Delta y^2)\right]\right\}_{m, n} \quad (3.36)$$

In Eq. (3.36), digital phase mask $\Phi(m, n)$ compensates the phase curvature introduced by MO.

CHAPTER 3 DIGITAL HOLOGRAPHY

As discussed in Section 3.3.1, in case the reference wave is a spherical wave and it does not have the same curvature with the object wave, compensation of the curvature difference is needed. In order to acquire correct phase image, curvature introduced by MO, curvature difference between reference and object waves and other phase aberrations need to be compensated. However, adjustment of these two quantities is a complex task which needs expertise and a priori knowledge of whole imaging system. Direct digital compensation methods, which the total phase aberrations is directly obtained from digital holograms and reconstructed images without the knowledge of propagation of reference wave and d_i and d_o of Eq. (3.35), have been explored[66]. Compensation of these two can be performed either in hologram plane or image plane or both planes. Different compensation methods will be introduced individually as follows:

Ferraro *et al.* reported digital compensation in the image plane [74]. Although it is based on in-line phase-shifting configuration with a planar reference wave, the method can also be extended to the off-axis configuration with spherical reference wave. First, the complex wavefront at the hologram plane is retrieved. Second, a correction phase factor from complex wavefront at hologram plane is found. Three ways are provided to find this correction phase factor. One way is to extract correction phase factor from a portion where circular fringes are not affected by the phase of specimen in phase image of complex wavefront in hologram plane and then extrapolate the fringes to the rest area of hologram. The 2nd way is to generate a synthetic phase distribution similar to the circular fringes by observation. The 3rd way is holographic interferometry— subtraction of hologram taken in area of specimen and hologram taken in nearby area of flat surface

without specimen. After the correction phase factor is found, numerical reconstructions for both the hologram and the correction wavefront are performed at the focusing distance—image plane. And subtraction of two reconstructed phase image at the focused plane is done to acquire correct phase image with compensation for inherent curvature.

Colomb *et al.* [75] proposed another phase compensation approach in the image plane. As the digital reference wave term is inside the Fresnel integral and digital phase mask is outside the Fresnel integral as in Eq. (3.36), usually they have to be adjusted individually. However, in this report, they merge the digital reference wave and the digital phase mask in a single entity which is outside the Fresnel integral of wavefront $O(x,y)$. This combined entity can be calculated by extraction from reconstructed phase values in image plane along line profiles in areas which are flat and serves as reference surfaces. The parameters of the merged single entity can be automatically adjusted by applying curve-fitting procedures on the extracted phase profiles. This approach compensates not only the quadratic phase curvature due to MO and the curvature difference between the reference and object waves at the same time in the image plane, but also other high-order aberrations introduced by other components of the set-up.

Colomb *et al.* reported one compensation method in hologram plane [76]. They record two holograms— one is hologram of specimen, the other is reference hologram with no specimen. The phase of the reference hologram is extracted out in hologram plane to compensate the phase curvature caused by MO, the curvature difference between the reference and object waves and other optical element in hologram plane. They also

demonstrate that in particular cases where the specimen does not have abrupt edges, the specimen's hologram itself can be used as the reference hologram.

Approaches that can place the compensation in the hologram plane or in the image plane have also been reported[78]. The compensation phase factor is defined by standard or Zernike polynomial models whose parameters are adjusted automatically by 2D fitting procedure applied to specimen areas known to be flat, rather than by 1D profile fitting in reference. It is also shown that this approach can not only compensate aberrations completely, but also can measure aberrations quantitatively, center the region of interest automatically and manually, perform numerical magnification.

Montfort *et al.* present a systematic and analytical study on the influence of different compensation positions (hologram or image plane) on the reconstructed image size and location in space [91]. In the 1st case the reference wave compensation is applied in the hologram plane and the curvature compensation related to MO is in the image plane; the 2nd case is a hologram plane approach in which a single compensation is applied in the hologram plane; the 3rd case is image plane approach where only one total compensation is applied in the image plane; the 4th case is a mixed approach where first-order phase correction is done in hologram plane and another compensation in image plane account for remaining high orders correction. The reconstruction distances of hologram plane approach, image plane approach and mixed approach are all different from the 1st ideal approach. The magnification and lateral location of reconstructed images of those three approaches compared to the ideal case are shown in table 3.2.

CHAPTER 3 DIGITAL HOLOGRAPHY

Approach	Y	Magnification	Shift
Ideal	$Y_{id}(x,y)$	1	(0, 0)
Hologram plane	$\exp[ik(d-d_{id})] \frac{1}{M_h} Y_{id}^I\left(\frac{x}{M_h}, \frac{y}{M_h}\right)$	$\frac{h_o-d}{h_o}$	(0, 0)
Image plane	$\exp[ik(d-d_{id})] \frac{1}{M_h} Y_{id}^I\left(\frac{x - \frac{d}{h_r} S_{Rx}}{M_i}, \frac{y - \frac{d}{h_r} S_{Ry}}{M_i}\right)$	$\frac{h_r-d}{h_r}$	$\left(\frac{d}{h_r} S_{Rx}, \frac{d}{h_r} S_{Ry}\right)$
Mixed	$\exp[ik(d-d_{id})] \frac{1}{M_i} Y_{id}^I\left(\frac{x}{M_i}, \frac{y}{M_i}\right)$	$\frac{h_r-d}{h_r}$	(0, 0)

Table 3.2 Summary of the Different Reconstructed Image Properties[91].

Table 3.2 presents the advantages of placing compensation in the hologram plane—tilt correction in the hologram plane allows an automatic centering of the region of interest in the image plane and the complete aberration compensation in the hologram plane is preserved for any reconstruction distance. Furthermore, phase aberrations at hologram plane will partially convert into image distortion during wave propagation. Compensation in hologram plane can avoid such an image distortion. The drawback of this solution is that the image is out of focus in the hologram plane. Therefore the areas used to determine the compensation are influenced by the diffraction pattern of the object. A minor adjustment in the image plane is needed which becomes a special mixed approach.

The image plane approach has the advantage of a focused image, and therefore it is easy to perform the phase compensation procedure without the influence of diffraction effect. The drawback is that the phase curvature is not compensated for all reconstruction distances and also the image is not centered in the reconstruction window. The image plane approach cannot avoid image distortion due to the propagation of phase

aberrations from hologram plane. The mixed approach has the same particularities as image plane approach except that the reconstructed image is centered.

In conclusion, the ability of digital holography to access complex wavefield digitally makes the numerical correction of the aberrations in the complex wavefront possible. Different approaches to digitally compensate phase deformations are introduced in this section. The correct complex wavefield of object can be obtained with these numerical compensation methods.

3.5 Review of Methods for Resolution Improvement in Digital Holography

In this section, we are going to provide a general review of the previous studies on the lateral resolution, its improvement and axial resolution and accuracy of DH.

3.5.1 Review of Lateral Resolution Analysis

Lateral resolution means resolution in x and y directions or in the plane of the object. Although DH has many advantages compared to conventional holography, its resolution is limited due to CCDs or other recording devices. Three factors contribute to this limitation, namely, the pixel averaging effect within the finite detection size of one pixel, a finite CCD aperture size limitation and the sampling effect due to a finite sampling interval. A review of these three factors is provided here.

Garcia-Sucerquia *et al.* [92] derived a criterion for lateral resolution, $\frac{\lambda}{2NA}$ where NA is the numerical aperture and showed that it is equals to $\frac{D}{(z^2+D^2)^{1/2}}$ where D is half CCD size and z is the distance from point source to CCD. This criterion is derived from numerical simulations. However how their simulations are done is not known. In their criterion, only finite CCD size is involved.

Jin *et al.* [93] discussed the influence of the pixel size, sampling interval and CCD size on the reconstructed image of the digital hologram. They concluded that the influence of the CCD pixel size determines the uniformity of brightness of the image. The influence of sampling interval causes a limitation on the object size to avoid overlapping of the reconstructed image. The influence of the CCD size restricts the resolution of the image. However, in the discussion of each of the three factors, the influences of the other two factors are ignored and not included. Therefore their results do not provide a comprehensive understanding of the influences of these factors in practice.

Kelly *et al.* [94] focused on two of the three factors: sampling effect and the effect of finite pixel size. In the sampling effect analysis, they pointed out that the sampling process creates an infinite number of replicas in the image. These replicas are separated from each other by a distance $f_s \lambda z$ where f_s is the sampling frequency. Each of the replicas is multiplied by a linear phase as well as some unimportant constant phase. Due to the pixel averaging effect, the intensity value recorded by the camera is weighted by its spatial frequency. And at spatial frequencies that are large relative to the pixel size, noise will become increasingly important in determining the accuracy of measurement.

All the above works [92-94] investigate the three factors individually. Interaction of them is the subject of some other researches.

Stern *et al.* [95] investigated the three factors. Their conclusions include that finite sampling interval leads to aliasing. The ways to overcome overlapping in different situations are discussed. The lateral resolution limit is $\lambda z/L_2$ in continuous reconstruction and is larger than $\max(\lambda z/L_2, \Delta x_2)$ where L_2 is the CCD size. Pixel averaging effect causes a degradation of the reconstructed image that is proportional to the effective detection size. This degradation can be modeled by an MTF if the Fresnel field is linearly encoded or is signal dependent if not linearly coded. However it investigates sampling and finite CCD size limitation effects together with the assumption of no pixel averaging effect and also the pixel averaging effect in the assumed condition of infinite CCD size but not the interaction of the three together. Therefore the arrived conclusions have limitations and are not comprehensive.

Picart *et al.* [96] discussed the influence of the transfer function of the discrete Fourier transform on the lateral resolution. Finite CCD size effect is included in the transfer function of the discrete Fourier transform. If only the finite CCD effect is considered that the lateral resolution is $\lambda z/(NP)$ where NP is the CCD size. The interaction of transfer function of the discrete Fourier transform and pixel averaging effect on the lateral resolution is also discussed. A criterion based on energy is used to judge the influence of pixel averaging effect on the intrinsic resolution $\lambda z/(NP)$. However, this criterion cannot tell how the two interact to determine the lateral resolution. An energy based criterion is not sufficient to indicate the resolution. Though some examples are

given, which factor plays the dominant role in different situations is not explicitly explained.

Kelly *et al.* [97] studied the above three factors and their interactions. In the individual study, it reported that sampling effect creates an infinite number of replicas in the image plane. Replicas are separated by a distance $\frac{\lambda z}{T}$ where T is sampling interval. Each replica is multiplied by a different linear phase as well as some unimportant constant phase factors. The effect of pixel averaging is to multiply the spatial frequency content of the signal by a *sinc* function. The effect of finite CCD size is to reduce the resolving ability of the DH system by convolving the product of the initial input and a quadratic phase term with a *sinc* function. The width of this *sinc* function is $\lambda z/(2D)$ where $2D$ is the CCD size. In the investigation of the interaction of the three factors, three regions are defined based on index $(D + \frac{\Delta}{2})/(\frac{\lambda z}{\xi})$ where Δ is object extent and 2ξ is sensing size of a single pixel. In the index smaller than 0.15, lateral resolution is determined by finite CCD size and sampling rate. In the range that index is larger than 0.15 but smaller than 3, both finite CCD size and pixel averaging effect contribute to the lateral resolution. In the range that index is larger than 3, lateral resolution is determined by pixel averaging effect. However, in the discussion of the interaction of factors, it only defines the domains dominated by finite camera extent or pixel averaging or both of them. The detailed relation of the two is not clearly unveiled. How they exactly interplay to determine resolution is still not known. Furthermore the boundaries of domains are arbitrarily defined and the percentage of domination is not provided. Therefore, further analysis is needed.

3.5.2 Review of Lateral Resolution Improvement

Many works on the improvement of lateral resolution have been reported. One direct method is to introduce MO into DH system [40, 50, 66, 75-81, 98-102]. The drawback of this method is the reduction of the field of view. And MO introduces unwanted curvature and other aberrations which need to be compensated to get correct result.

Another method to enhance lateral resolution is to use aperture synthesis method which improves the lateral resolution and does not reduce the field of view. It can be categorized into three approaches. One approach translates the CCD position and records multiple holograms at different positions to collect more object information at larger diffraction angle. The second approach provides multiple illumination directions with fixed CCD position to record multiple holograms. Different diffraction angles of object information can be projected onto and recorded by the CCD with multiple illumination directions. In the third method the specimen is rotated with a fixed CCD position and illuminating angle. Different diffraction angles of the object are recorded on to the CCD. Larger diffraction angles correspond to higher object frequencies and larger numerical aperture. Thus better lateral resolution can be expected. After recording, information recorded by different holograms needs to be integrated by aperture synthesis.

Reference [103-111] illustrated the first aperture synthesis approach of translation of CCD positions. Kreis *et al.* [103] introduced the concept of synthesizing a larger aperture by two CCD arrays. Theoretical analysis and simulations of this concept are performed. An improvement in resolution can be obtained and the resolution relates to

the size of the synthesized aperture. It is also pointed out that the depth of focus is reduced at the same time.

Massig *et al.* [104] reported the achievement of a synthetic aperture in an off-axis arrangement. The aperture is constructed from nine holograms with different CCD recording positions with overlapping area between adjacent holograms. In the synthesis, one hologram serves as reference. Relative positions of other holograms are found by the magnitudes of the cross correlation between the concerned hologram and the reference hologram. A Fourier holographic setup is used. A comparison of image resolution reconstructed from single hologram and from a synthesis of nine holograms is shown in Fig. 3.11. An increase of the resolution by a factor of 2.5 was presented.

Figure 3.11 (a) reconstruction of the object from a single hologram; (b) reconstruction of the object from the synthesis of nine holograms [104].

Martinez-Leon *et al.* [109] reported a system of single-exposure on-line digital holography with improved resolution using a synthetic aperture. The setup is based on Mach-Zehnder interferometer with spherical reference and object beams of the same curvature. This setup can provide a 3.5x magnification. Nine holograms with three rows

and three columns are recorded and synthesized into one hologram with a correlation algorithm. The synthesized hologram is 2700×2400 pixels in size compared to a single hologram (1368×1024 pixels). The vertical resolution can be improved from G5E4 to G6E4 (in the USAF target) which corresponds to 90.51 line pairs/mm as seen in Fig 3.12. The lateral horizontal resolution was improved to G6E1. It is mentioned that the tilts between different holograms need to be compensated. To quantify the resolution enhancement, correlation method is utilized with the image of high resolution hologram as a reference.

Figure 3.12 Reconstruction of a USAF resolution target from (a) single hologram; (b) synthesized hologram and (c) a high resolution digital hologram [109].

Di *et al.* [108] reported an approach based on synthetic aperture by using linear CCD scanning to obtain digital holographic images with high resolution and a wide field of view. Digital lensless Fourier transform hologram with a large area of $3.5\text{cm} \times 3.5\text{cm}$ (5000×5000 pixels) was obtained. The numerical reconstruction image shows that a theoretically minimum resolvable distance of $2.57\mu\text{m}$ and a field of 4mm view can be achieved at a distance 14cm at 632.8nm wavelength as seen in Fig 3.13.

Figure 3.13 (a) the reconstructed holographic image of a synthetic aperture of $3.5\text{cm} \times 3.5\text{cm}$ (5000×5000 pixels), (b) group 25 in (a) and (c) a partial of (b) in square [108].

Gyimesi *et al.* [107] addressed two problems. The limited resolution of CCD restricts the field of view and the finite size of CCD restricts the lateral resolution in the reconstructed image. To improve the lateral resolution of the reconstructed image, synthesis of aperture is used. With a 10×10 synthetic aperture of 4096×4096 pixels, an improvement of resolution by a factor of 6 is achieved as in Fig. 3.14. On the other hand the limited pixel size of CCD can be overcome by demagnification the hologram before the CCD. However demagnification of hologram automatically decreases lateral resolution. Trade off is always needed between lateral resolution and field of view.

Figure 3.14 (a) the reconstructed holographic image of original aperture of 3.8mm×3.8mm (512×512 pixels), (b) the reconstructed holographic image of synthetic aperture of 30.1mm×30.3mm (4096×4096 pixels) [107].

Claus *et al.* [106] reported the realization of synthetic aperture method by moving CCD with a motorized x - y stage. In this way a larger numerical aperture is obtained. This enables a more detailed reconstruction. The enhanced resolution of about three times can be seen in Fig. 3.15. In the experiment a lensless Fourier holography setup is used. Larger numerical aperture also enables a smaller depth of field. The depth of field can be increased by applying the extended depth of field method.

Figure 3.15 (a) the reconstructed image of original aperture of 3000×3000 pixels, (b) the reconstructed holographic image of synthetic aperture 8000×8000 pixels [106].

Mico *et al.* reported a synthetic aperture method by shifting CCD to different off-axis positions based on a modified Gabor-like holography configuration [111]. Different spatial-frequency content from different CCD positions is merged into one by aperture synthesis. A resolution gain close to 2 is achieved as seen in Fig. 3.16.

Figure 3.16 (a) the image without aperture synthesis (horizontal and vertical NA are 0.043 and 0.032 respectively); (b) the image with aperture synthesis (horizontal and vertical NA are 0.086 and 0.064 respectively) [111].

References [82, 112-135] illustrated the second aperture synthesis approach by multiple illumination directions. The multiple of illumination directions can be carried out by translation of a single light source, tilted illumination directions, utilizing multiple differently located light sources and by use of gratings.

Reference [112, 113, 116, 117] reported an aperture synthesis approach by the translation of a single light sources. References [112, 113, 116] are based on Fourier holography configuration while reference [117] are based on Fresnel holography configuration. But the principles are similar. We use reference [112] to illustrate this approach. Alexandrov *et al.* [112] reported synthetic aperture optical microscopy in which high resolution images are obtained from the synthesis of a set of Fourier holograms. The illumination wave is a planar wave and has polar angle θ_i and azimuthal angle ϕ_i as seen in Fig 3.17 (a). This illumination wave corresponds to the spatial frequencies $\gamma_\xi = \frac{\sin\theta_i \cos\phi_i}{\lambda}$ and $\gamma_\eta = \frac{\sin\theta_i \sin\phi_i}{\lambda}$ in the object plane. The object spatial frequency ν that diffracts or scatters into the angle α_{out} as shown in Fig 3.17 (b) can be written as $\nu = \frac{|\sin\theta_i - \sin\alpha_{out}|}{\lambda}$. If the CCD is fixed at $\alpha_{out} = 0$ and has size of $-\frac{L}{2} \leq x \leq \frac{L}{2}$, the range of object spatial frequencies captured by the CCD corresponds to the following range of sample spatial frequencies:

$$\nu_{\xi,min/max} = \mp \frac{L}{2\lambda d} + \gamma_\xi$$

$$\nu_{\eta,min/max} = \mp \frac{H}{2\lambda d} + \gamma_\eta$$

By changing the illumination azimuthal angle ϕ_i from 0° , 90° , 180° , 270° , γ_ξ and γ_η changes correspondingly as in Fig 3.17 (c). Therefore, different object spatial frequency range ν_ξ and ν_η can be covered. By synthesis of all the reconstructed images from

individual holograms, lateral resolution enhanced image can be acquired. Fig 3.18 (a) shows the phase image reconstructed from a single hologram with $\phi_i = 0$ and $\theta_i = 49^\circ$. Fig 3.18 (b) and (c) present the phase images of the selected section of the image synthesized by 4 reconstructed images. The corresponding synthetic NA is 0.76. A grating of 1200 lines/mm can be fully resolved. Fig 3.18 (d) shows a confocal image with NA equal to 1. It can be seen that the image quality reconstructed by aperture synthesis in Fig. 3.18 (c) is better than that of confocal microscope in Fig. 3.18 (d).

Fig 3.17 (a) Orientation of object and recording planes, and illumination and reference waves with respect to the optical axis z ; (b) depiction of the incident and scattered or diffracted waves in the plane of incidence, and the collection solid angle; (c) angle of spatial frequencies covered in four separate holographic recordings (0° , 90° , 180° , 270°). The range covered by a single 0.75NA lens shown [112].

Fig 3.18 images of a reflection grating with resolution of 1200 lines/mm: (a) reconstructed image from azimuthal angle 0° ; (b), (c) phase image of selected and area of synthesized image; (d) confocal microscope image [112].

Brueck's Group reported works on imaging interferometric microscopy which is an aperture synthesis approach based on the tilted illumination [128-132]. The main idea is tilted (off-axis) illumination, which shifts the higher object frequencies into the band pass of the objective lens as shown in Fig. 3.19 (a) [130]. A collimated illumination beam is incident on the object at an angle of incident β that is greater than the collection angle of the objective. A zero-order beam is brought around the objective lens with an auxiliary optical system and re-introduced on the low NA side to interfere with diffracted beams transmitted through the objective. By appropriately adjusting the divergence, direction and phase of the zero-order beam, the interference shifts the collected diffracted information back to high frequency. By changing the angle of β and the corresponding zero-order beam, different frequencies of the object are recorded and

synthesized to obtain an extension of the frequency space coverage of the object and therefore improve the resolution of the image. With this approach and using 633 wavelength, images of non-periodic 180nm pattern ($\frac{\lambda}{3.52}$) and of a 170nm grating ($\frac{\lambda}{3.72}$) is achieved [130]. A related concept, imaging interferometric lithography has been introduced for lithographic image formation [128]. A major advantage for microscopy is that the partial images can be electronically manipulated, whereas in the lithography case the image information is chemically stored in the photoresist and is not directly accessible. A simple case with only two offset partial images, one each in orthogonal directions is presented and demonstrates the possibility of resolving 0.5 μm features using a 0.4 NA objective with 633nm wavelength ($\frac{\lambda}{1.2}$) [129]. Base on reference [130], a new approach to achieve the same resolution using a simpler and more robust configuration is proposed [131]. Instead of introducing a zero order on the low NA side, a reference wave on the front side of the objective is used. Two ways are discussed. The first is to add an off-axis illumination beam at an incident angle close to the edge of the imaging system NA in front of the object as illustrated in Fig. 3.19 (b). The other is the injection of a reference beam into the objective using a beam splitter between the object and the objective lens as illustrated in Fig. 3.19 (c). The advantage of this approach is that reduced or no access to the image pupil plane is required. This makes the implementation on existing microscope much easier. The recorded images are at low frequencies which reduces demands on the resolution of CCD. The system is more stable than conventional interferometric microscopy.

Figure 3.19 (a) Optical arrangement for imaging interferometric microscopy [130]. $\alpha = \sin^{-1}(NA)$, β is the incident beam angle of incidence, and β_{ref} is the angle of the reference beam onto the image plane; (b) structured illumination with an off-axis illumination beam at an incident angle close to the edge of the imaging system NA in front of the object[131]; (c) structured illumination by the injection of a reference beam into the objective using a beam splitter between the object and the objective lens[131].

A group including Zalevsky, Mico and Granero *et al.* made great contributions in aperture synthesis method by changing the illumination angles [82, 118-127, 133]. The underlying principle of all these approaches is to illuminate the object with a set of tilted beams.

Reference [114] reported a resolving approach for off-axis digital holographic microscopy where a microscope objective is used. In this approach the single illumination point source shifts in sequential mode and holograms are recorded at each shift position. Holograms recorded at different illumination positions are superimposed. Each shift of the illumination beam generates a shift in the object spectrum in such a

way that different spatial-frequency bands are transmitted through the objective lens as seen in Fig. 3.20. The lateral resolution was enhanced by a factor 3 times in x and y direction and 2.4 in the oblique directions. And finally reconstruction with a lateral resolution $1.74\mu\text{m}$ in Fig 3.21 was demonstrated. The set-up is a phase step digital holographic setup. Reference [82] also reported an approach by shifting a single illumination point source to record additional frequency bands for aperture synthesis. But, instead using an off- axis digital holographic microscopy, a common-path phase-shifting digital holographic microscopy configuration is adopted. Reference [124] reported an aperture synthesis method by shifting a single illumination point source to provide different tilted illuminations too. In this reference, a very simple digital in-line holographic microscopy configuration is used.

Figure 3.20 Fourier transform of the addition of different recorded holograms [114].

CHAPTER 3 DIGITAL HOLOGRAPHY

Figure 3.21 (a) Image obtained with 0.1 NA lens and conventional illumination. (b) Supperesolved image obtained with the synthetic aperture. The G9E2 corresponding to the resolution limit using the proposed method is marked with arrow [114].

In references [119-121, 127], different illumination angles, on-axis illumination and off-axis illumination, are generated by gratings. The grating is located between the laser source and object. The different diffraction orders of the grating are used to generate different illumination angles. The on-axis illumination and off-axis illumination are recorded so that higher frequencies of the object spectrum can be collected to generate synthetic aperture, thus improving the resolution. Differences between these references [119-121, 127] are the setup configuration and the algorithms to recover additional frequency bands. References [123, 135] place the grating after the object in such a way that high order diffracted components are redirected towards the imaging device. As seen in Fig. 3.22 (a), a diffraction grating is inserted between the object and CCD. Without the grating, only the central portion will reach the CCD area [seen in Fig. 3.22 (b)]. With the grating, the zero order of the grating does not affect the propagation of the different spectral portions [Fig. 3.22 (c)] but the grating diffracts additional spatial-frequency portions towards the CCD aperture [seen in Fig. 3.22 (d)]. Since this new spectral portions reach the CCD obliquely, it will be possible to recover separately each one of them because they will not overlap at the Fourier domain. Once again, this separation depends on the properly selection of the diffraction grating.

References [118, 125] reported an aperture synthesis approach based on a 2D array of mutually incoherent vertical-cavity self-emitting laser sources. Instead of changing the angle of one illumination light, they use incoherent illumination points array to generate and record multiple holograms. Each of these holograms records different spatial frequency bands with different illumination sources. In reconstruction, each frequency band is adjusted for focusing. After focusing, the super-resolved object is reconstructed from the synthesized hologram by Fourier transformation. In reference [118], different sources are switched on sequentially. Therefore holograms containing different frequency bands are recorded sequentially. The arrangement of the different frequency bands is performed digitally at a later stage. However in [125], all sources illuminate the object simultaneously. Aperture synthesis is performed by a single CCD snapshot instead of multiple sequentially recorded holograms.

Reference [126] reported a single-exposure super resolution imaging in digital holographic microscopy configuration. It combined angular multiplexing with wavelength multiplexing. Red, blue and green beams simultaneously illuminate the sample with different tilted angles 0 , θ_B and θ_G as seen in Fig. 3.23. Three different color images containing different spatial frequency content of the input sample are simultaneously directed towards CCD. A color CCD records the hologram. In reconstruction, the information included in each color channel is retrieved by analyzing the three RGB channels of CCD. Based on reference [126], reference [133] presented a different experiment configuration where the color CCD is replaced by a monochrome CCD. To maintain the single-exposure character, the object field of view is restricted and hologram recording is based on image-plane wavelength dispersion spatial multiplexing to record three band-pass images by a single CCD snapshot.

Figure 3.23 Experimental setup for single-exposure super-resolved interferometric microscopy [126].

Yuan *et al.* [115] reported an aperture-synthesis approach based on a pulsed digital holography system. A single laser pulse from the light source is amplitude divided into two trains with specially designed incident angles for both object and reference beams. In the object beams, a single laser pulse is divided into three sub-pulses with a given time delay and different incident angles. Three sub-pulsed reference beams with the

same time delay of different tilt angles interfere with object diffraction fields at the CCD plane successively. The different tilt angles guarantee the carrier frequencies of the three sub-holograms are different so that their Fourier spectra will separate from each other in spectrum. The same time delay between the three pairs of recording beams is adjusted longer than the pulsed width of the laser source to ensure an incoherent overlapping of the three sub-holograms. This ensures the successively recorded holograms can be incoherently superimposed into one hologram. Therefore the diffraction fields with different spatial frequency range of object can be recorded in one hologram. In reconstruction, the respectively reconstructed images of the sub-holograms are synthesized. An improvement of 1.78 times of the lateral resolution of the synthesized image can be obtained as shown in Fig. 3.24.

Figure 3.24(a)-(c) Intensity images of three different object illumination angle; (d) aperture synthesized image; (a')-(d') the central parts enclosed in the rectangles of (a)-(d) [115].

The use of 2D grating was first reported by Paturzo *et al.*[134]. A 2D hexagonal phase diffraction grating G is inserted between the object and the CCD as seen in Fig. 3.25 (a). In conventional holography, all the rays scattered by the object propagate forward to the CCD plane. But only the central ray can be collected by the CCD due to the finite size

of CCD as shown in Fig. 3.25 (b). When the grating is inserted, six further rays can reach CCD as depicted in Fig. 3.25 (c). Each of the six rays is produced by the first diffraction orders of the grating. As can be seen in Fig. 3.25 (c), the CCD aperture is augmented up to 3 times along each of the three directions at 120° thanks to the hexagonal geometry. The reconstruction is performed by two steps—a first propagation from hologram to grating plane and then multiplication of the grating function and another propagation from grating plane to image plane.

Figure 3.25 (a) Scheme of DH recording setup: Fourier configuration in off-axis mode; (b) without the grating in the setup and (c) with the grating in the setup [134].

References [136, 137] illustrated the third aperture synthesis approach by specimen rotation. Binet *et al.* [136] reported an aperture synthesis method by rotating the object relative to the CCD with unchanged illumination light direction. At different rotation angle, different object spatial frequencies are projected onto the CCD and recorded. In the synthesis processes of different holograms, the relative position and phase of each

hologram has to be given. The overlap of holograms enables the estimation and compensation of their relative positions and phase with a speckle cross-correlation algorithm. The overlap size is half of the CCD size. By synthesis of 33 holograms with 128×128 pixels into a 2048×128 pixels hologram, the resolution gain is 16 in horizontal direction as shown in Fig 3.21. Zhang *et al.* [137] also reported similar method as in [136] but with a Mach-Zehnder setup.

Figure 3.21 Image reconstructed from (a) a single hologram of 128×128 pixels (b) synthetic aperture composed of 33 holograms merged coherently. It is of 2048×128 pixels [136].

Of the three approaches of aperture synthesis, the first one by translation of CCD has a limitation on the object spatial frequency recorded by hologram. When CCD is away from the optical axis, the light corresponding to higher object spatial frequency is collected with larger diffraction angle by the CCD. If the incident of reference wave is not changed at different CCD position, sampling interval sets a limitation of the maximum frequency to avoid spectrum overlapping. The other two approaches do not suffer from this limitation as the angles of the light projected on CCD are not affected by the object spatial frequency. In the second method, object spatial frequency recorded

by CCD is controlled by the illumination light angle. Larger illumination angles diffract higher object frequencies onto the CCD. In the third method, the object spatial frequency recorded by CCD is controlled by rotation angle of the object. Larger object rotation angle diffracts higher object frequencies onto the CCD. In the third method, as the object position and phase are changed during the rotation, additional work to compensate the position and phase change must be needed.

A system with both high lateral resolution and large FOV is always desired. The second and third methods leave no way for field of view extension. Compared to a single hologram, these two methods only add additional object spatial frequencies by synthesis. However the first method, not only includes additional object spatial frequencies, but also includes new object information at the new position. Therefore it permits enhancement of both lateral resolution and FOV at the same time. Until now, only few [107, 108] have reported enhancement of both lateral resolution and FOV at the same time. Nakatsuji *et al.* [138] reported a synthetic aperture method used only for extension of field of view but with lateral resolution unchanged.

In [107], the lateral resolution is improved by hologram stitching. But the FOV is extended by demagnification of hologram size which automatically decreases lateral resolution. Trade off is always needed between lateral resolution and FOV. Di *et al.* [108] reports improvement of both lateral resolution and FOV at the same time by hologram stitching, using Fourier holography configuration. The main limitation of Fourier holography is that it can only reconstruct image at one plane. In contrast, the more general Fresnel holography geometry does not suffer from such a limitation. The

improvement of both lateral resolution and FOV at the same with Fresnel holography has not been studied.

3.5.3 Review of Axial Resolution

Axial resolution means resolution in z direction or out-of-plane direction of the object. The analysis and enhancement of lateral resolution have been reviewed in the last two sections. As a 3D measurement technique, like the lateral resolution, the axial resolution is also a very important parameter to define the system performance in axial measurement.

However there is no such criterion to define axial resolution as Rayleigh criterion for lateral resolution. The quantization effect due to the camera bit depth in the process of analog-to-digital conversion (ADC) of the hologram intensity has been considered. In digital holography technique, the axial dimension is measured by optical path length difference (OPD). OPD is derived from the phase of the measured complex object wavefront. The relation between OPD and phase is, $OPD = \lambda \frac{Phase}{2\pi}$. As the OPD information represented in the phase of the object wave is encoded by the reference wave into the interference pattern—hologram, the quantization involving rounding or truncating of the values of hologram causes phase and therefore OPD errors.

If the ADC quantizes analogs signal into an eight bit data, there is $2^8=256$ discrete quantization levels. The quantization effect gives phase resolution of $\frac{2\pi}{256}$ which corresponds to OPD resolution of $\frac{\lambda}{256}$. The relationship between the height of object and the OPD is also related to the DH mode of recording. In reflection mode, the measured

CHAPTER 3 DIGITAL HOLOGRAPHY

height is related to OPD as $Height = \frac{OPD}{2}$. For transmission DH, the measured height is related to OPD as $Height = \frac{OPD}{n-n_0}$ where n is the refractive index of the object and n_0 is the refractive index of the background. This OPD resolution $\frac{\lambda}{256}$ provides an axial resolution of $\frac{\lambda}{256 \times (n-1)}$ in transmission mode and axial resolution of $\frac{\lambda}{256 \times 2}$ in reflection mode. However, beside quantization effect, other effects of digital holography are also involved to affect the axial resolution. This has been manifested by reported works that the axial resolution of $\frac{\lambda}{256 \times (n-1)}$ in transmission mode ($n_0 = 1$) and of $\frac{\lambda}{256 \times 2}$ in reflection mode cannot be achieved in practice. References associated with axial measurement accuracy or resolution have been reported as in Table 3.3.

Axial measurement errors can be one of the reasons in case they are much larger than the theoretical resolution and overwhelm it. In such case the above theoretical cannot be achieved.

Author/ Reference	Axial Resolution	Resolution	DH Mode
Cuche, E [67]	height resolution	better than 10nm	reflection mode
Marquet, P [40]	phase accuracy	2 – 4°	transmission mode
Colomb, T [76]	phase noise	2.7°	transmission mode
Charriere, F [98]	phase error	4.1°	either
Rappaz, B [79]	thickness accuracy	1µm	transmission mode
Carl, D [83]	OPD resolution	30nm	either
Colomb, T [139]	height resolution	better than 2 nm	reflection mode

Mann, C. J [140]	Phase noise (thickness variance)	10°/30nm	transmission mode
------------------	-------------------------------------	----------	-------------------

Table 3.3 Reported results associated with axial measurement resolution and accuracy.

Therefore the factors which cause the axial measurement errors and thus influence the axial measurement performance needs to be indentified and their interactions in the influence need to be analyzed. However, few works have been reported on this. There is only a discussion of shift variant properties of DH image [97] on axial measurement in which the existence of space-dependent phase in the point spread functions of two point sources is focused. However the DH system is considered as an ideal system with no pixel averaging effect within single pixel area and thus makes this discussion not comprehensive enough. Therefore there is a lot of room for investigation of the factors which cause the axial measurement errors and the interactions between them.

3.6 Possible Exploration Directions

Based on the review in section 3.5, three directions can be further explored.

First is the analysis of lateral resolution. Three factors contribute to the DH lateral resolution limitation, namely, the pixel averaging effect within the finite detection size of one pixel, a finite CCD aperture size limitation, and the sampling effect due to a finite sampling interval. A review of these three factors contributing to the DH lateral resolution limitation has been done [52, 92-97, 141, 142]. Most of these works investigate the three factors individually [92-94, 141]. Some discuss the interaction of two of them [95, 96, 142]. Only reference [97] studied the above three factors

interactively. However this study is not sufficient in aspects discussed in section 3.5.1. Therefore, further analysis of lateral resolution is needed.

Second is the simultaneous improvement of both lateral resolution and field of view. System with both high lateral resolution and FOV is always desired. But from the review in section 3.5.2, it can be seen that though there are many works done on lateral resolution enhancement [103-110, 112-118, 136, 137, 143], few permit simultaneous enhancement of both lateral resolution and field of view [107, 108]. However, as discussed in section 3.5.2 these methods suffer certain limitations. Therefore, further exploration in this direction is meaningful.

Third is the analysis of axial measurement errors. As discussed in section 3.5.3, there are no criteria for axial measurement resolution. Though the CCD quantization effect is considered to provide a resolution limitation, such resolution is hard to achieve in practice. One cause of this is the axial measurement errors which may be much larger than the resolution set by the quantization effect. But until now, few work have been done on the analysis of factors which causes axial measurement errors, let alone thorough investigation of the interaction between them and the weightages of different factors on the measurement errors. Therefore such a work is necessary and meaningful.

SUMMARY

In this chapter, the review of DH technique is provided which can be concluded as below:

1. Different configurations of DH are summarized.

2. Digital recording process is introduced. The object spectrum collected by DH system depends on the CCD size, object extent and distance from hologram plane to CCD plane. To guarantee correct reconstruction, the object spectrum needs to satisfy the requirement of sampling effect of CCD in the recording process. Furthermore, in off-axis geometry, the tilt angle of reference wave should be large enough to separate the zero order and first order in spectrum in recording.
3. Numerical reconstruction process is illustrated. First step is the hologram pre-processing at the hologram plane where the object wave is extracted from hologram and the tilt induced by off-axis geometry is compensated numerically. After that, in the numerical propagation from hologram plane to image plane, Fresnel and convolution methods are introduced and compared.
4. Aberration compensation methods for phase imaging are reviewed. In order to obtain correct quantitative phase information, phase aberrations due to MO, unequal curvatures of the reference wave and object wave, imperfections of optics and other issues should be compensated. Different compensation methods are reviewed here.
5. Research works related to DH resolution are reviewed. As resolution is one of the most important indices of DH and also is its main limitation compared to traditional holography, investigation and improvement of the resolution of DH technique is necessary and meaningful. Therefore, in this part of work, research works on the analysis of lateral resolution, lateral resolution improvement and analysis of axial resolution and accuracy are reviewed.
6. Based on the review of DH technique, possible exploration directions are discussed and pointed out.

CHAPTER 4 LATERAL RESOLUTION ANALYSIS OF DIGITAL HOLOGRAPHY

4.1 Introduction

Although digital holography (DH) has many advantages compared to conventional holography, its lateral resolution is limited due to CCDs or other recording device parameters. As discussed in last chapter, three factors contribute to this limitation, namely, the pixel averaging effect due to the finite pixel size, a finite CCD aperture size, and the sampling effect due to a finite sampling interval.

In this chapter, the influences of the three factors on DH lateral resolution are investigated. The lateral resolution is determined by the interaction of finite CCD size and pixel averaging effect. A 3D map presenting their relationship is provided. The lateral resolution can be determined by the values of these two factors. The domains dominated by them are explained along with their accuracy. As DH system is proven to be space variant, object extent also influences the lateral resolution. The influences of the object extent on lateral resolution are discussed. Sampling effect which puts a requirement on object extent to avoid aliasing and guarantee the lateral resolution is also investigated. From this study and for a fixed system geometry, the lateral resolution capability of the DH system can be readily estimated. This analysis can be used for the determination of system lateral resolution and its improvement. The resolution performance of both in-line and off-axis systems is studied and an example for resolution determination in a practical system is provided.

This chapter is organized as follows. In section 4.2, the DH system with pixel averaging, finite CCD size and sampling effect is expressed. Effects of the interaction between pixel averaging and finite CCD size on the image resolution are investigated in section 4.3. Firstly the effects on point spread function (PSF) (or impulse response) of a DH system are analyzed, followed by a discussion on an object with finite extent. The requirements on object size and its influences on the lateral resolution are discussed. The limitation due to the sampling effect is presented in section 4.4. Examples of the above limitations on the resolution capability of in-line and off-axis geometries are given in section 4.5. Section 4.6 summarizes the main conclusions of this chapter.

4.2 Holography Expression with Finite CCD Size, Pixel Averaging and Sampling Effect

Digital holography includes two major aspects: digital recording and numerical reconstruction. We consider that the coding process with reference wave and the decoding process with the conjugate reference wave are exactly inverse of each other so that they can be ignored in the analysis of a DH system. With Fresnel diffraction integral[45], including the finite CCD size, pixel averaging and sampling effect, the reconstructed wavefront $Rf(x)$ of the object wavefield $f(x)$ can be expressed as:

$$Rf(x_2) = \text{Fresnel} \left\{ \left[f(x) \otimes \exp\left(\frac{j\pi}{\lambda z} x^2\right) \right] \times \text{rect}\left(\frac{x}{2D}\right) \right\} \otimes \left\{ \text{rect}\left(\frac{x}{2p}\right) \times \sum_{-\infty}^{+\infty} \delta(x - nT) \right\} \quad (4.1)$$

where $\otimes$ denotes convolution, λ is wavelength used in recording, z is distance between the object plane and CCD plane, $2D$ is size of CCD sensing chip, $2p$ is the size of a single pixel and T is the spatial sampling interval. The ability of a CCD to resolve spatial frequency at CCD plane is determined by sampling interval T [95, 97]. Smaller T provides better lateral resolution capability. However, T cannot be made very small due to two reasons. Firstly, T cannot be

manufactured as small as desired due to the limits in current manufacture capability. Secondly, as pixel size is physically equal or smaller than the sampling interval, smaller sampling interval means smaller pixel size too. Furthermore smaller pixel size increases noise level. In order to balance noise level, sampling interval should not be too small. We use x_2 to present the coordinate in image plane and x to present the coordinate on the CCD plane. Function $rect\left(\frac{x}{2D}\right)$ and $rect\left(\frac{x}{2p}\right)$ are rectangle functions with width $2D$ and $2p$ respectively. The ratio $\frac{2p}{T}$ is the fill factor. The *Fresnel* in Eq. (4.2) defines the Fresnel transform as

$$Fresnel(k(x)) = \exp\left(-\frac{j\pi}{\lambda z} x_2^2\right) \mathcal{F}\{k(x) \times \exp\left(-\frac{j\pi}{\lambda z} x^2\right)\}_{f=-\frac{x_2}{\lambda z}} \quad (4.2)$$

Due to the separable property of Fresnel transform, only x dimensional is considered and the two-dimensional extension can be readily deduced. In Eq. (4.1), the convolution $f(x) \otimes \exp\left(\frac{j\pi}{\lambda z} x^2\right)$ expresses the process of object wave propagation from the object plane to the CCD plane. The finite size of CCD presents in the term $rect\left(\frac{x}{2D}\right)$. The second convolution with function $rect\left(\frac{x}{2p}\right)$ accounts for the pixel averaging effect over the entire CCD sensing area. Finally the term $\sum_{-\infty}^{+\infty} \delta(x - nT)$ is the CCD sampling effect. By substituting Eq. (4.2) into Eq. (4.1) we get

$$Rf(x_2) = \exp\left(-\frac{j\pi}{\lambda z} x_2^2\right) \times \mathcal{F}\left\{\left\{\left[f(x) \otimes \exp\left(\frac{j\pi}{\lambda z} x^2\right)\right] \times rect\left(\frac{x}{2D}\right)\right\} \otimes rect\left(\frac{x}{2p}\right) \times \sum_{-\infty}^{+\infty} \delta(x - nT) \times \exp\left(-\frac{j\pi}{\lambda z} x^2\right)\right\}_{f=-\frac{x_2}{\lambda z}} \quad (4.3)$$

where $\mathcal{F}$ denotes Fourier transform.

According to the property of Fourier transform and convolution, Eq. (4.3) can be written as

$$Rf(x_2) = R_0f(x_2) \otimes \sum_{n=-\infty}^{\infty} \delta(x_2 - \frac{n\lambda z}{T}) \quad (4.4)$$

where

$$R_0f(x_2) = \exp\left(-\frac{j\pi}{\lambda z} x_2^2\right) \mathcal{F} \left\{ \left\{ \left[f(x) \otimes \exp\left(\frac{j\pi}{\lambda z} x^2\right) \right] \times \text{rect}\left(\frac{x}{2D}\right) \right\} \otimes \text{rect}\left(\frac{x}{2p}\right) \right\} \times \exp\left(-\frac{j\pi}{\lambda z} x^2\right) \right\}_{f=-\frac{x_2}{\lambda z}} \quad (4.5)$$

Eq. (4.5) shows that the finite CCD size and pixel averaging effect interact together. And the sampling, as in Eq. (4.4), acts on this interaction by generating multiple replicas of $R_0f(x_2)$ with an interval $\frac{\lambda z}{T}$. Therefore, the interaction of the finite CCD size and pixel averaging can be investigated first, followed by the investigation of the sampling effect on the interaction.

4.3 Finite CCD Size and Pixel Averaging Effect

In this section, interactions of finite CCD size and pixel averaging on the resolution are investigated. The point spread function (PSF) of DH system influenced by these two factors is investigated first, followed by the analysis of an object with finite extent.

4.3.1 Point Spread Function (PSF) Analysis of DH

4.3.3.1 Investigation and Results

We consider a point object $f(x) = \delta(x)$ and investigate the PSF of DH system with Eq. (4.5). In such a case, the object wavefront at the CCD is $\delta(x) \otimes \exp\left(\frac{j\pi}{\lambda z} x^2\right) = \exp\left(\frac{j\pi}{\lambda z} x^2\right)$. In Eq. (4.5), the parameter $2p$ in $\text{rect}\left(\frac{x}{2p}\right)$ is the contribution of the pixel averaging and the parameter $\frac{\lambda z}{D}$ represents the finite CCD size effect as per the finite chirp function properties. Fig. 4.1 plots the

width of PSF for different values of $2p$ and $\frac{\lambda z}{D}$ where according to the Rayleigh criterion, resolution equals to half width of PSF. The 3D surface shows the interaction between the pixel size and CCD size on resolution of the reconstructed field. The resolution can be indexed with different values of $2p$ and $\frac{\lambda z}{D}$. To illustrate this in further detail, three rows of this 3D surface corresponding to parameter $\frac{\lambda z}{D}$ equals to constant values of 2.906×10^{-6} , 3.7781×10^{-5} , 7.4981×10^{-5} along with the corresponding derivatives of the width with respect to $x = 2p$ are plotted in Fig. 4.2. Three columns corresponding to $2p$ equal to 2.906×10^{-6} , 3.7781×10^{-5} , 7.4981×10^{-5} along with the corresponding derivatives of the width with respect to $y = \frac{\lambda z}{D}$ are shown in Fig. 4.4.

Figure 4.1 Width of PSF of DH system (the width is defined by the distance of the first two zeros of the PSF as in Fig. 4.6 (h) and 4.7(h)) with different values of $2p$ and $\frac{\lambda z}{D}$ to show the interaction of these two parameters on the resolution. x axis represents $2p$ and y axis represents $\frac{\lambda z}{D}$. Both axis contain 255 points with the value from 1.1625×10^{-6} to 1.488×10^{-4} with an interval of 0.58125×10^{-6} .

Figure 4.2 (a), (c) and (e) are the 4th, 64th, 128th rows of Fig. 4.1 and correspond to $\frac{\lambda z}{D} = 2.906 \times 10^{-6}$, 3.7781×10^{-5} , 7.4981×10^{-5} respectively; (b) (d) (f) are the derivatives with respect to $x = 2p$ of profile (a), (c) and (e) respectively. x axis represents the normalized $2p$ by the corresponding $\frac{\lambda z}{D}$. The fluctuations in (d) and (f) are due to insufficient padding in reconstruction. With sufficient padding, the curves can be smooth. But the trend of (d) and (f) can be seen.

Figure 4.3 Difference of the width of PSF and $2p$ of all the rows in Fig. 4.1 (calculation: $\frac{\text{width} - 2p}{2p}$).

It can be seen from Fig. 4.2 that the width of PSF increases as $2p$ increases while the derivative oscillates before settling to unity as $\frac{2p}{(\lambda z)/D}$ becomes larger. And the width of PSF increases as $\frac{\lambda z}{D}$ increases and once again the derivative oscillates and approaches unity as $\frac{(\lambda z)/D}{2p}$ becomes larger or $\frac{2p}{(\lambda z)/D}$ becomes smaller in Fig. 4.4. Therefore the ratio $\frac{2p}{(\lambda z)/D}$ is an important index to indicate the weight of the two factors on resolution. To further understand this interaction, we plot the relationship between $\frac{width-2p}{2p}$ v/s $\frac{2p}{(\lambda z)/D}$ in Fig. 4.3 and the relationship between $\frac{width-(\lambda z)/D}{(\lambda z)/D}$ v/s $\frac{2p}{(\lambda z)/D}$ in Fig. 4.5 where width is the PSF width. As shown in Fig. 4.3, when the ratio $\frac{2p}{(\lambda z)/D}$ is equal to 6, 30, 60 and 120, the difference is 10%, 2%, 1% and 0.5% respectively. In Fig. 4.5, when the ratio $\frac{(\lambda z)/D}{2p}$ is equal to 2, 6 and 27, the difference is 10%, 1% and 0.0018%. When the difference between the PSF width and $2p$ is less than 10%, the resolution is mainly determined the pixel size and similarly when the difference between the PSF width and $\frac{\lambda z}{D}$ is less than 10%, then the resolution is determined by the CCD size. Alternately, when $\frac{2p}{(\lambda z)/D}$ is greater than 6, the pixel size decides the resolution and when $\frac{2p}{(\lambda z)/D}$ is smaller than 0.5, then the resolution is mainly determined by CCD size. For the case when $\frac{2p}{(\lambda z)/D}$ is in the region $0.5 < \frac{2p}{(\lambda z)/D} < 6$, both of the two effects contribute to the determination of the lateral resolution.

Figure 4.4 (a), (c) and (e) are the 4th, 64th, 128th columns of Fig. 4.1 and corresponds to $2p = 2.906 \times 10^{-6}$, 3.7781×10^{-5} , 7.4981×10^{-5} respectively; (b) (d) (f) are the according derivatives with respect to $y = \frac{\lambda z}{D}$ of profile (a), (c) and (e) respectively. x axis represents the normalized $\frac{\lambda z}{D}$ by the corresponding $2p$. The fluctuations in (b), (d) and (f) come from insufficient padding in reconstruction. With sufficient padding, the curves can be smooth. But the trend of (b), (d) and (f) can be seen here already.

Figure 4.5 Difference of the width of PSF and $\frac{\lambda z}{D}$ of all the rows in Fig. 4.1 (calculation: $\frac{\text{width} - \frac{\lambda z}{D}}{\frac{\lambda z}{D}}$)

4.3.3.2 Analysis of Results

To further clarify the results in the previous section, each step of Eq. (4.5) is examined in detail as shown in Fig. 4.6 and 4.7. The column on the left lists the functions of each step in the spatial domain while the one on the right shows the corresponding function in the Fourier domain. F1 to F5 denote different functions in Fig. 4.6 and 4.7. In step 1, F1 convolutes with F2. The spectrum of finite chirp function F1 is as shown in Fig. 4.6 (b) and Fig. 4.7 (b) with bandwidth $\frac{2D}{\lambda z}$. The amplitude is almost flat and is equal to one within this bandwidth and zero outside [45]. The spectrum of the rectangle function F2 is $\text{sinc}(2pf)$ of bandwidth $\frac{1}{p}$ (distance between the first two zeros). The convolution in spatial domain corresponds to a multiplication in spectral domain. Therefore the bandwidth of F3 is determined by the smaller of $\frac{1}{p}$ and $\frac{2D}{\lambda z}$. The amplitude spectrum of F3 is of the form $\text{sinc}(2pf) \times \text{rect}(\frac{\lambda z}{2D}f)$ and the corresponding complex amplitude in space is $\text{sinc}(\frac{2p}{\lambda z}x) \text{rect}(\frac{x}{2D}) \exp(j\frac{\pi}{\lambda z}x^2)$. After multiplication with F4 (the finite CCD size $\text{rect}(\frac{x}{2D})$ is taken into consideration in numerical reconstruction, but not shown in Eq. (4.5) to simplify the expression), the complex amplitude of F5 is

$\text{sinc}\left(\frac{2p}{\lambda z}x\right)\text{rect}\left(\frac{x}{2D}\right)$. When $\frac{2p}{\lambda z/D} \geq 1$ as in Fig. 4.6, the distance between the two first zeros of the $\text{sinc}\left(\frac{2p}{\lambda z}x\right)$ is $\frac{\lambda z}{p}$ which is smaller than $2D$. Therefore the signal $\text{sinc}\left(\frac{2p}{\lambda z}x\right)\text{rect}\left(\frac{x}{2D}\right)$ is a truncated $\text{sinc}\left(\frac{2p}{\lambda z}x\right)$ as in Fig. 4.6(g) and its Fourier transform is a rectangle rather than a sinc function shape as seen in Fig. 4.6(h). However when $\frac{2p}{\lambda z/D} < 1$, as in Fig. 4.7, $\text{sinc}\left(\frac{2p}{\lambda z}x\right)\text{rect}\left(\frac{x}{2D}\right)$ takes the form of $\text{rect}\left(\frac{x}{2D}\right)$ modulated by $\text{sinc}\left(\frac{2p}{\lambda z}x\right)$ in amplitude Fig. 4.7(g) and its Fourier transform is more of a sinc function rather than a rectangle shape as seen in Fig. 4.7(h).

From the results and analysis, it can be concluded that PSF of DH system falls into two categories:

- (a) $\frac{2p}{\lambda z/D} \geq 1$: PSF follows $\text{rect}\left(\frac{x}{2p}\right)$ function. As the ratio $\frac{2p}{\lambda z/D}$ becomes large, its shape approaches $\text{rect}\left(\frac{x}{2p}\right)$ and its width approaches $2p$. When the ratio is larger than 6, the width of PSF is mainly determined by $2p$ with more than 90% weightage.
- (b) $\frac{2p}{\lambda z/D} < 1$: PSF follows $\text{sinc}\left(\frac{2D}{\lambda z}x\right)$ function. As the ratio $\frac{2p}{\lambda z/D}$ become small, its shape approaches $\text{sinc}\left(\frac{2D}{\lambda z}x\right)$ and its width approaches $\frac{\lambda z}{D}$. When the ratio is smaller than 0.5, the width of PSF is mainly determined by $\frac{\lambda z}{D}$ with more than 90% weightage.

 Figure 4.6 PSF investigation of DH system as in Eq. (4.5) in case of $\frac{2p}{\lambda z/D} \geq 1$.

Figure 4.7 PSF investigation of DH system as in Eq. (4.5) in case of $\frac{2p}{\lambda z/D} < 1$.

4.3.2 Object with Finite Extent

When object $f(x)$ is not a point source but has a finite spatial extent L_o , it can be considered as a sum of point sources along extent L_o . Its spectrum at the CCD plane is the sum of all the individual spectrum of these points. We take three points A, B and C to analyze their spectra at

the CCD plane (Fig. 4.8 (a)). The wavefronts from A, B, C at the CCD plane are $\exp[j\frac{\pi}{\lambda z}(x + x_i)^2] \times \text{rect}(\frac{x}{2D})$ ($i = A, B$ and C). The bandwidth of A, B and C extend $(-\frac{D}{\lambda z} + \frac{x_i}{\lambda z}, \frac{D}{\lambda z} + \frac{x_i}{\lambda z})$ according to finite chirp function properties. If A and C are at the boundaries of the object such that the object is centered at $(x_A = -\frac{L_o}{2}, x_C = \frac{L_o}{2}$ and $x_B = 0$), they contribute to the highest frequency of the object signal at the CCD as shown in Fig. 4.8 (b), (c) and (d). The bandwidth of object $f(x)$ with extent L_o is $(-\frac{D}{\lambda z} - \frac{L_o}{2\lambda z}, \frac{D}{\lambda z} + \frac{L_o}{2\lambda z})$.

Figure 4.8 Spectra of points along object with extent L_o at the CCD plane. (a) Positions of the three points: B is at the center, A and C are the edges. (b), (c) and (d) are the spectra of point A, B and C respectively at the CCD plane.

Consider edge point C as an example to study the influence of the object size on the reconstructed image resolution, as in Fig. 4.10 and 4.11. The reconstructed image of C is

$$\exp[j\frac{\pi}{\lambda z}(\frac{L_o}{2})^2] \times \mathcal{F}\{\text{sinc}[\frac{2D}{\lambda z}(x - \frac{L_o}{2})] \times \text{rect}(\frac{x}{2D}) \times \exp(-j2\pi \frac{L_o}{2\lambda z}x)\}_{f_x = -\frac{x_2}{\lambda z}}.$$

For an arbitrary point at $x = x_0$ ($|x_0| \leq L_o/2$), the reconstructed image is

$$\exp\left[\frac{j\pi}{\lambda z}(x_0)^2\right] \times \mathcal{F}\left\{\text{sinc}\left[\frac{2p}{\lambda z}(x - x_0)\right] \times \text{rect}\left(\frac{x}{2D}\right) \times \exp\left(-j2\pi\frac{x_0}{\lambda z}x\right)\right\}_{f_x=-\frac{x_2}{\lambda z}} \quad (4.6)$$

The PSF $\mathcal{F}\left\{\text{sinc}\left(\frac{2p}{\lambda z}x\right) \times \text{rect}\left(\frac{x}{2D}\right)\right\}_{f_x=-\frac{x_2}{\lambda z}}$ in the previous section is just a special case of Eq.

(4.6) when $x_0 = 0$. If we expand the Fourier transform of Eq. (4.6), we have

$$\exp\left[\frac{j\pi}{\lambda z}(x_0)^2\right] \times \left\{ \left[\exp\left(j\frac{\pi x_0}{p}x_2\right) \times \text{rect}\left(\frac{x_2}{2p}\right) \right] \otimes \text{sinc}\left(\frac{2D}{\lambda z}x_2\right) \otimes \delta(x_2 - x_0) \right\} \quad (4.7)$$

Eq. (4.7) indicates that DH is a space variant system. The factor $\exp\left[\frac{j\pi}{\lambda z}(x_0)^2\right]$ is spatially dependent. The convolution $\left[\exp\left(j\frac{\pi x_0}{p}x_2\right) \times \text{rect}\left(\frac{x_2}{2p}\right) \right] \otimes \text{sinc}\left(\frac{2D}{\lambda z}x_2\right)$ is also spatially dependent due to phase factor $\exp\left(j\frac{\pi x_0}{p}x_2\right)$ as seen in Fig. 4.9. Therefore, the reconstructed images of point sources from different parts of the object have different amplitude profiles though parameters $2p$ and $\frac{\lambda z}{D}$ are same. At some positions, such as $x_0 = 0.0174\text{m}$ in Fig. 4.9(b), information of this point is almost lost. Hence the size of object also influences the DH resolution performance.

Figure 4.9 (a) Amplitude of convolution $\left[\exp\left(j\frac{\pi x_0}{p}x_2\right) \text{rect}\left(\frac{x_2}{2p}\right) \right] \otimes \text{sinc}\left(\frac{2D}{\lambda z}x_2\right)$ changes as x_0 changes. The y -axis represents x_0 changing from 0 to 0.032m and x - z plane presents the profile of amplitude profiles of $\left[\exp\left(j\frac{\pi x_0}{p}x_2\right) \text{rect}\left(\frac{x_2}{2p}\right) \right] \otimes \text{sinc}\left(\frac{2D}{\lambda z}x_2\right)$ at according x_0 . (b) Profiles when x_0 is 0, 0.0062, 0.0126, 0.0174m in (a).

In order to investigate the influence of object size on resolution, each step of Eq. (4.5) for point C is examined in details as shown in Fig. 4.10 and 4.11. When $2p \geq \frac{\lambda z}{D}$ as in Fig. 4.10, the reconstructed amplitude image follows the $\text{rect}\left(\frac{x-L_o}{2p}\right)$ function as in Fig. 4.10 (h). To ensure that the main information is not lost for all point sources along extent L_o , it should be satisfied that the information of edge points A and C is not lost. This requires that the main lobe or the first zeros of $\text{sinc}(2pf)$ in Fig. 4.10 (d) lies within the bandwidth $\text{rect}\left[\frac{\lambda z}{2D}\left(f + \frac{L_o}{2\lambda z}\right)\right]$ in Fig. 4.10 (b) i.e.

$$\frac{D}{\lambda z} - \frac{L_o}{2\lambda z} \geq \frac{1}{2p} \quad (4.8)$$

If we further require that $2p$ mainly determines the resolution of the image along the whole object extent L_o , according to the previous section, at least the sixth zero of the $\text{sinc}(2pf)$ function lies within the bandwidth of $\text{rect}\left[\frac{\lambda z}{2D}\left(f + \frac{L_o}{2\lambda z}\right)\right]$ i.e.,

$$\frac{D}{\lambda z} - \frac{L_o}{2\lambda z} \geq \frac{6}{2p} \quad (4.9)$$

In the case of $2p < \frac{\lambda z}{D}$ as in Fig. 4.11, the reconstructed image follows $\text{sinc}\left[\frac{2D}{\lambda z}(x - x_0)\right]$ function (Fig. 4.11 (h)). In order to make sure that the main information is not lost along extent L_o , the bandwidth of $\text{rect}\left[\frac{\lambda z}{2D}\left(f + \frac{L_o}{2\lambda z}\right)\right]$ in Fig. 4.11 (b) should be within the main lobe of $\text{sinc}(2pf)$ in Fig. 4.11 (d). Fig. 4.11 (f) shows the product of $\text{sinc}(2pf) \times \text{rect}\left(\frac{\lambda z}{2D}\left(f + \frac{L_o}{2\lambda z}\right)\right)$.

This requires:

$$\frac{D}{\lambda z} + \frac{L_o}{2\lambda z} < \frac{1}{2p} \quad (4.10)$$

If $\frac{\lambda z}{D}$ mainly determines the resolution of the image along the whole object extent L_o , the

bandwidth of $rect[\frac{\lambda z}{2D}(f + \frac{L_o}{2\lambda z})]$ has to lie within the half width of the main lobe of $sinc(2pf)$:

$$\frac{D}{\lambda z} + \frac{L_o}{2\lambda z} < \frac{1}{2} \times \frac{1}{2p} \quad (4.11)$$

Figure 4.10 Investigation of flow chart in Fig. 4.1 for the case of $\frac{2p}{(\lambda z)/D} \geq 1$ with object of finite extent L_o .

 Figure 4.11 Investigation of flow chart in Fig. 4.1 for the case of $\frac{2p}{(\lambda z)/D} < 1$ with object of finite extent L_o .

Figure 4.12 Examples of the violation of Eq. (4.8). (a) shows the spectra of two point sources at the CCD plane where point P2 follows the condition of Eq. (4.8) and point P1 does not; (b) shows the reconstructed images of the points P1 and P2. (c) and (d) are the magnified images of P1, P2 in (b) to show details.

Figure 4.13 Examples of violation of Eq. (4.10). (a) the spectra of three point sources at the CCD plane where point P2 follows the condition of Eq. (4.10) and points P1 and P3 violate Eq. (4.10); (b) the reconstructed images of the three points. (c), (d) and (e) are the magnified images of P1, P2 and P3 in (b).

Fig. 4.12 gives an example of the violation of Eq. (4.8). Point P2 follows Eq. (4.8) whereas points P1 does not. The intensity of image P2 is much higher than that of P1 due to the lower amplitude of side lobes of $\text{sinc}(2pf)$ where the spectrum of P1 lies (Fig.4.12 (a)). A single point source at P1 is imaged as two points which is not expected for a good imaging system. An example of the violation of Eq. (4.10) is shown in Fig 4.13. Point P2 satisfies the condition of Eq. (4.10) whereas points P1 and P3 do not. The intensity of image P2 is much higher than those of P1 and P3 due to the lower amplitude of side lobes of $\text{sinc}(2pf)$ too. The distances of the first zeros in (c), (d) and (e) are 3.9153×10^{-5} , 3.6735×10^{-5} and 5.2508×10^{-5} m respectively which suggests that violation of Eq. (4.10) results in poor resolution. Though both P1 and P3 violate Eq. (4.11), their shapes are different and depend on where their bandwidths lie within $\text{sinc}(2pf)$ envelope. Hence violation of Eq. (4.8) and (4.10) induces not only lower energy images but also position-dependent instability and poor resolution. To avoid these effects, Eq. (4.8) and (4.10) should be followed.

Fig. 4.9 is the case of $2p < \frac{\lambda z}{D}$. Eq. (4.10) and (4.11) correspond to $0.008 \leq |x_0| < 0.016$ and $|x_0| < 0.008$ respectively. When $|x_0| < 0.008$, Fig. 4.9 (a) shows that PSF width is in the range: $\frac{\lambda z}{D}$ to $1.1 \times \frac{\lambda z}{D}$. When $0.008 \leq |x_0| < 0.016$, the width of main lobe becomes gradually larger but does not exceed $\frac{\lambda z}{D} + 2p$. However, when $|x_0| \geq 0.016$, it is difficult to distinguish the main lobe from the side lobes. And the width of reconstructed signal has greater uncertainty as in Fig. 4.9 (b) when $x_0=0.0174\text{m}$. Therefore, conditions of Eq. (4.8) and (4.10) should be followed to avoid information loss and Eq. (4.9) and (4.11) are further recommended for consistent resolution in the field.

4.4 Sampling Effect

The product with $\sum_{-\infty}^{+\infty} \delta(x - nT)$ in Eq. (4.1) in the spatial domain corresponds to a convolution with $\sum_{-\infty}^{+\infty} \delta(f - n/T)$ in the frequency domain. This requires that the extent of signal spectrum should be in the range $-\frac{1}{2T}$ to $\frac{1}{2T}$. Thus, to avoid spectrum overlap, the following should be satisfied

$$\left(\frac{D}{\lambda z} + \left|\frac{L_o}{\lambda z}\right|\right) / \frac{1}{T} \leq \frac{1}{2} \quad (4.12)$$

where L_o is the object extent. Eq. (4.12) exerts an extra condition for object extent L_o and CCD size $2D$ to follow.

4.5 Examples of System Analysis

Consider the following recording parameters

- pixel size: $2p=4.65\mu m$
- CCD size: 1280×960 with 100% fill factor of CCD pixel
- wavelength $\lambda = 633nm$

which gives

- $\frac{D_x}{\lambda z} = \frac{1280 \times 4.65 \times 10^{-6}}{2 \times 633 \times 10^{-9} \times z} = \frac{4.701 \times 10^3}{z}$
- $\frac{1}{2p} = \frac{1}{4.65 \times 10^{-6}} = 215.05 \times 10^3 l/mm$

4.5.1 In-line Geometry

For $z=100mm$, we have $2p/\frac{\lambda z}{D_x} \cong 0.22$ which is smaller than 0.5. Therefore, the maximum resolvable spatial frequency of this DH system is determined by $\frac{\lambda z}{D_x} = 21.27\mu m$ with an accuracy of 98%. The minimal resolvable object size for this DH system indexed from Fig. 4.1

or deduced from Fig. 4.5 is $10.85\mu m$. This resolution can only be achieved at the center of image plane due to the spatial variant property of DH. The minimal resolvable object size R considering the position according to Eq. (4.10) and (4.11) is

$$\begin{array}{ll} 10.85\mu m < R \leq 11.7\mu m & 0 < |x_0| \leq 3.83mm \\ 11.7\mu m < R \leq 12.92\mu m & 3.83mm < |x_0| \leq 10.637mm \\ R \text{ is unstable and poor} & |x_0| > 10.637mm \end{array}$$

where x_0 is the location from the center of the image plane. At the same time Eq. (4.12) requires that object extent should follow $L_{ox} < 13.587mm$. Thus $|x_0|$ should not exceed $6.79mm$.

4.5.2 Off-axis Geometry

In off-axis geometry, the bandwidth of the zero order is twice the extent of the first order in the spectrum. Thus the highest frequency present in the signal has to satisfy

$$\frac{D}{\lambda z} + \frac{L_o}{2\lambda z} \leq \frac{1}{8T} \quad (4.13)$$

As $2p$ is normally smaller than T , Eq. (4.11) is valid and the resolution is governed by $\frac{\lambda z}{D}$ with an accuracy of more than 99%.

4.6 Conclusion

The main conclusions of this work are briefly summarized in this section.

The lateral resolution of a DH system has been studied in previous works [92, 93, 95-97]. References [92, 93, 95, 96] reported a lateral resolution of $\frac{\lambda z}{D}$ with the assumption that only the finite CCD effect is considered in a DH system. The influence of pixel averaging effect was discussed in references [93-95]. The intensity recorded by the CCD is weighted by its spatial

frequency by function $\text{sinc}(2pf)$ where $2p$ is the detection size of a single pixel [93, 94]. At spatial frequencies that are larger than the first zeros of $\text{sinc}(2pf)$, noise will become increasingly important in determining the accuracy of measurement. Reference [95] also reported that pixel averaging effect causes spectrum attenuation by pixel MTF $\text{sinc}(2pf)$. These results agree with the discussion of function $\text{sinc}(2pf)$ in Fig. 4.7 and 4.11 in the case of $2p/(\frac{\lambda z}{D}) < 1$ of this work. However, the interaction of pixel averaging effect and finite CCD size on the lateral resolution was not considered in these previous studies [93-95]. Reference [96] discussed the interaction of CCD size and pixel averaging on the lateral resolution. A criterion based on energy is used to judge the influence of pixel averaging effect on the lateral resolution $\lambda z/(2D)$ determined by the finite CCD size. However, this criterion cannot tell how the two interact to determine the lateral resolution. An energy based criterion is not sufficient to indicate the lateral resolution. Reference [97] investigated the interaction of CCD size and pixel averaging based on an index $(D + \frac{\Delta}{2})/(\frac{\lambda z}{p})$ where Δ is object extent. If the index is smaller than 0.15, the lateral resolution is determined by the finite CCD size. For the range between 0.15 and 3, both the finite CCD size and pixel averaging effect contribute to the lateral resolution. For index values greater than 3, the lateral resolution is determined by the pixel averaging effect. But the interplay between these variables and its effect on the lateral resolution is still not known. Furthermore the boundaries of domains are arbitrarily defined and the percentage of domination is not provided.

In this chapter, the specific interaction of CCD size and pixel averaging effect on the lateral resolution is analyzed. The lateral resolution of a DH system is determined by the interaction of parameters $2p$ and $\frac{\lambda z}{D}$. The extent to which the lateral resolution is dominated by $2p$ and $\frac{\lambda z}{D}$ is

determined by the ratio $2p/\frac{D}{\lambda z}$. The relation curves of the domination and the ratio $2p/\frac{D}{\lambda z}$ are also presented. Therefore, for any given values of $2p$ and $\frac{\lambda z}{D}$, the lateral resolution can be easily obtained with the help of these curves.

Though the lateral resolution is determined by $2p$ and $\frac{\lambda z}{D}$, the lateral resolution away from center becomes worse due to the space variant property of DH system. This property has been noticed in reference [97]. But the relationship between the spatial location and the lateral resolution was not investigated. In this chapter, this relationship is investigated. When $2p < \frac{\lambda z}{D}$, the position-wise minimal resolvable object size R is:

$R_{best} < R \leq 1.1 \times \frac{\lambda z}{2D}$	$\frac{x_o}{2\lambda z} \leq \frac{1}{4p} - \frac{D}{\lambda z}$
$1.1 \times \frac{\lambda z}{2D} < R \leq \frac{1}{2} \times (\frac{\lambda z}{D} + 2p)$	$\frac{1}{4p} - \frac{D}{\lambda z} < \frac{x_o}{2\lambda z} \leq \frac{1}{2p} - \frac{D}{\lambda z}$
R is poor due to information loss	$\frac{x_o}{2\lambda z} > \frac{1}{2p} - \frac{D}{\lambda z}$

When $\frac{\lambda z}{D} \leq 2p$, the position-wised minimal resolvable object size R is

$R_{best} < R \leq 1.1 \times p$	$\frac{x_o}{2\lambda z} \leq \frac{D}{\lambda z} - \frac{3}{p}$
$1.1 \times p < R \leq \frac{1}{2} \times (\frac{\lambda z}{D} + 2p)$	$\frac{D}{\lambda z} - \frac{3}{p} < \frac{x_o}{2\lambda z} \leq \frac{D}{\lambda z} - \frac{1}{2p}$
R is poor due to information loss	$\frac{x_o}{2\lambda z} > \frac{D}{\lambda z} - \frac{1}{2p}$

where R_{best} denotes the lateral resolution at the center of image plane.

The sampling effect has been analyzed in references [94, 95]. The sampling process creates an infinite number of replicas in the image. These replicas are separated from each other by a

distance $\frac{\lambda z}{T}$ in space and each of the replicas is multiplied by a linear phase factor as well as some unimportant constant factor. Therefore the object size should satisfy $L_o \leq \frac{\lambda z}{T}$ [94, 95]. But, beside the replicas in the spatial domain, the effect of sampling in spectrum domain also needs to be considered. By consideration of the sampling effect in both spatial and spectral domains, to avoid aliasing, we find that the extent of object L_o should satisfy the condition $(\frac{D}{\lambda z} + \left| \frac{L_o}{\lambda z} \right|) / \frac{1}{T} \leq \frac{1}{2}$ which is in agreement with an earlier study [97].

The point spread function (PSF) is used in the analysis in this work. The concept of PSF for coherent and incoherent optical systems was introduced and explained by Goodman [45]. The application of PSF for holography system analysis and especially for the resolution analysis has been reported by Kreis [110, 144, 145] and by Christoph [141]. There are some differences between these prior works and the work in this chapter.

The first difference is the system module. Only the CCD size and sampling effect were considered in past works [110, 141, 144, 145]. In this section, additionally, the pixel averaging and spatial variance property for extended objects are also considered. This system module is much closer to practical system and also much more complex.

Secondly, previous studies [95-97, 110, 144, 145] simplified the system expression by the expansions, reorganization and simplification using mathematical theorems and laws such that they can be easily understood and analyzed. During this process, assumptions are made such as setting some complex terms to their extreme values e.g simplifying $\text{sinc}(\Delta \xi x') \text{comb}(\Delta \xi x')$ to $\delta(x')$ in references [144, 145]. As a result, some constraints are ignored and only special cases are discussed. However, instead of simplifying the system, the aim of this work is to investigate

the relationship between different system constraints and the lateral resolution of the complex system. In such case, the simplification of the system expression is not necessary and assumptions do not have to be made. Therefore the complex system expression is kept for a comprehensive investigation of the system. In the investigation, by linearly changing the parameter in concern, the width of the point spread function is calculated and plotted with complex system expression. Based on curve fitting, the relation between the parameter in concern and the lateral resolution is quantitatively determined.

Finally since no assumptions are made and no factors are set to their extreme values, the result is valid for general situations instead of only for special cases as in previous studies. Furthermore, the result of this study is presented in a quantitative manner. However, the results of the earlier studies are the simplified forms of equations, e.g. Eq. (18) in references [144, 145] or Eq. (12) in reference [110].

CHAPTER 5 LATERAL RESOLUTION IMPROVEMENT BY APERTURE SYNTHESIS

5.1 Introduction

Compared to conventional holography, digital holography (DH) has many advantages including access to quantitative amplitude and phase information [42-44]. However, the lateral resolution of DH is limited by the digital recording device and its field of view (FOV) is restricted by the limited resolution of pixelated detectors. A system with both high lateral resolution and large FOV is always desired. Microscope objectives can be used to improve the lateral resolution [71, 72]. But they decrease the FOV at the same time. Factors affecting the lateral resolution have been investigated [52, 93, 97, 146]. Hologram stitching is one aperture synthesis method that has been researched to overcome the lateral resolution limit [103-106, 109, 112-117, 136, 137, 143]. But its capability for both lateral resolution improvement and FOV enlargement has only been reported in limited references [107, 108]. In reference [107], while the lateral resolution is improved by hologram stitching, the FOV is extended by demagnification of hologram size which automatically decreases the lateral resolution. Tradeoff between the lateral resolution and FOV is needed [107]. Reference [108] reports the improvement of both the lateral resolution and FOV at the same time by hologram stitching, using Fourier transform holography. The main limitation of Fourier holography is that it can only reconstruct image at one plane. In contrast, the more general Fresnel holography geometry does not suffer this limitation. To our knowledge, the simultaneous improvement of both lateral resolution and FOV with Fresnel holography has not been reported.

In this chapter, the ability of hologram stitching for simultaneous improvement of both the lateral resolution and FOV is demonstrated with a more general Fresnel holography setup. The impact of aperture synthesis on the lateral resolution is investigated theoretically at first. In the experiment, the synthesis is executed by moving the compact digital holographic system in two directions. Nine holograms are recorded and stitched into one hologram.

The lensless Fresnel holography geometry used in this chapter has been shown to provide lensless magnification [54, 55, 88]. In this chapter, its capability to provide a larger NA (numerical aperture) and therefore better lateral resolution is demonstrated. By using two diverging beams with the same divergence at CCD, the object spectrum is compressed by a factor related to the magnification as compared to the geometry using two collimated beams.

5.2 Theoretical Analysis

5.2.1 Off-axis Fresnel DH Microscope

The lensless DH microscope system used in this work is shown in Fig. 5.1. The diverging incident beam is divided into two beams by a beam splitter. The object beam illuminates the sample and the reference beam is incident on a plane mirror. The reflected light from the sample interferes with that from the mirror at the CCD plane. The divergences of the two beams are matched at the CCD plane to produce straight line interference fringes. The geometry of the setup determines the magnification of this system. The magnified view of sample is obtained by reconstruction with a plane reference wave at reconstruction distance

$$Z = \frac{z_r z_s}{z_r - z_s} \quad (5.1)$$

from the hologram [45, 65]. The lateral magnification Mag of the reconstructed image is [65]

$$Mag = \frac{z_r}{z_r - z_s} \quad (5.2)$$

where z_r is the distance between light source and the CCD plane and z_s is the distance from object to CCD plane, Fig. 5.2. This system presents an off-axis reflection microscopic geometry via a simple and compact optical setup. This setup is best suited for highly specular micro-size objects. The distance from the point source to object is constrained by the beam splitter and therefore restricts the magnification of the system [54].

Figure 5.1 Setup of off-axis lensless Fresnel DH microscope. (BS: beam splitter.)

Fig. 5.2 shows the object beam (AB) originates at point A and interacts with an arbitrary point (C) located at x_o on the object plane. The object beam illuminates point C at an angle θ_1 . For $|x_o| \ll z_r - z_s$ (which is mostly the case in DH), we have

$$\theta_1 \cong \sin \theta_1 \cong tg \theta_1 = \frac{x_o}{z_r - z_s} = -\frac{|x_o|}{z_r - z_s} \quad (5.3)$$

The spatial frequencies of the object diffract the light with the undiffracted light propagating at the same angle θ_1 . The zero order impinges on CCD, a distance l_1 away from optical axis z can be expressed as in Eq. (5.4):

$$l_1 = z_r \times tg \theta_1 = \frac{|x_o| z_r}{z_r - z_s}. \quad (5.4)$$

Higher frequencies diffract away from this zero order with larger angles. If the frequency f_i diffracts at angle θ_i from optical axis z , then we have

$$\frac{\theta_i - \theta_1}{\lambda} \cong \frac{\sin \theta_i - \sin \theta_1}{\lambda} = f_i \quad (5.5)$$

Since the CCD has a finite size, $2D$, only light at angles between $-\frac{D - |x_o|}{z_s}$ and $\frac{D + |x_o|}{z_s}$ can be collected. From Eq. (5.5), this angle range corresponds to object frequencies in the range from

$-\frac{D - |x_o| z_r}{\lambda z_s}$ to $\frac{D + |x_o| z_r}{\lambda z_s}$. Therefore the bandwidth of object frequencies collected by the CCD is

$$\text{Bandwidth} = \frac{D + \frac{|x_o| z_r}{z_r - z_s}}{\lambda z_s} - \left(-\frac{D - \frac{|x_o| z_r}{z_r - z_s}}{\lambda z_s} \right) = \frac{2D}{\lambda z_s} \quad (5.6)$$

Eq. (5.6) indicates the bandwidth of arbitrary point x_o arriving at CCD is independent of position x_o . The lateral resolution of the system depends on bandwidth. The smallest detail that can be resolved by this system is $R = \frac{\lambda z_s}{2D}$.

Figure 5.2 Schematic of off-axis Fresnel DH geometry. $l_1, l_2, l_3, l_4, s, D, z_r$ and z_s are all physical length (of positive values). x_o denotes the position of point C in the coordinate.

The off-axis reference wave originating at point R interferes with the object wave at the CCD as in Fig. 5.2. In this geometry, the carrier frequency f_c is determined by angle θ_2 as

$$f_c = \frac{\sin\theta_2}{\lambda} \cong \frac{s}{z_r} \quad (5.7)$$

Object frequency bandwidth will be modulated by the carrier frequency. The spread of the interference pattern about the carrier frequency is from f_l (minimum) to f_r (maximum) which are determined by angle θ_3 and θ_4 respectively as

$$f_l = \frac{\sin\theta_3}{\lambda} \cong \frac{l_2}{z_r} \quad (5.8)$$

$$f_r = \frac{\sin\theta_4}{\lambda} \cong \frac{l_3+s}{z_r} \quad (5.9)$$

where l_2 and l_3 , according to geometric relationship in Fig. 5.2, are equal to

$$l_2 = s - l_4 = s - \frac{D-l_1}{z_s} (z_r - z_s) = s - \frac{D - \frac{|x_o| z_r}{z_r - z_s}}{z_s} (z_r - z_s) \quad (5.10)$$

$$l_3 = \frac{D+l_1}{z_s} (z_r - z_s) = \frac{D + \frac{|x_o| z_r}{z_r - z_s}}{z_s} (z_r - z_s) \quad (5.11)$$

Therefore the bandwidth of the interference pattern is

$$Bandwidth_{InterferencePattern} = f_r - f_l = \frac{(z_r - z_s)}{z_r} \frac{2D}{\lambda z_s} = \frac{1}{Mag} \frac{2D}{\lambda z_s} \quad (5.12)$$

Interference of object beam and reference beam can be considered as signal encoding. From the comparison of Eq. (5.6) and Eq. (5.12), it can be seen that the bandwidth after encoding by reference beam is compressed by factor equal to the magnification (Mag). Hence, the requirement on CCD sampling interval is reduced. Though the object bandwidth is compressed in the encoding process, object information recorded is not lost and remains the same. Therefore, when the object is reconstructed with a planar wave at distance z of Eq. (5.1), the lateral resolution $\frac{\lambda z_s}{2D}$ can be expressed with reconstruction distance z as

$$R = \frac{\lambda z_s}{2D} = \frac{\lambda z}{2D} \frac{z_r - z_s}{z_r} = \frac{1}{Mag} \frac{\lambda z}{2D} \quad (5.13)$$

Some may consider this system to be equivalent to the geometry in which both object illumination and reference beams are collimated waves and object located at z from CCD plane. But that is not the case, since this system provides a NA of $Mag \times \frac{D}{z}$ while the collimated beam geometry provides a NA of $\frac{D}{z}$. Therefore the lateral resolution of our system is Mag times of the lateral resolution in the collimated beam geometry, as given by Eq. (5.13).

5.2.2 Aperture Synthesis

The lateral resolution of DH system is limited by pixel averaging effect, finite CCD size, sampling effect and the object extent. The interaction of these factors on the lateral resolution are investigated and presented in the previous chapter with the assumption that the coding process (recording) and the decoding process (reconstruction) are exactly inverse of each other. Accordingly, for an off-axis setup with plane reference and object waves, $\frac{\lambda z}{2D}$ determines the lateral resolution where $2D$, λ and z are the CCD aperture size, wavelength of light wave and distance between the object and CCD respectively. Though the lateral resolution depends on position, it cannot be worse than $1.1 \times \frac{\lambda z}{2D}$ within the allowed object size [146]. Therefore lateral resolution can be improved by increasing CCD aperture size, $2D$. Hologram stitching is one way to do this.

In this chapter, the coding and decoding processes are taken into consideration for a more precise analysis of the effects of hologram stitching on the lateral resolution in the off-axis DH system. The reconstructed image $R_0 f(x_2)$ of the object $f(x)$ is

$$R_0 f(x_2) = \mathcal{F}^{-1} \{ \mathcal{F} \left\{ \left[f(x) \otimes \exp \left(j \frac{\pi}{\lambda z} x^2 \right) \right] \times \text{rect} \left(\frac{x}{2D} \right) \right\} \times \text{sinc}[(2p(f_x + f_c))] \times G(f_x) \} \quad (5.14)$$

where $\mathcal{F}$ denotes a Fourier transform, $\otimes$ denotes convolution, $2p$ is the pixel size and G is transfer function of free space propagation. Without loss of generality a one-dimensional analysis is studied due to the separable property of the Fresnel transform. The Fresnel approximation gives $G(f_x)$ as

$$G(f_x) = \exp[-j\pi\lambda d f_x^2] \quad (5.15)$$

The point spread function (PSF) is obtained when $f(x) = \delta(x)$ where $\delta(x)$ is a delta function.

Thus we get

$$PSF = \mathcal{F}^{-1} \{ \mathcal{F} \left\{ \exp \left(j \frac{\pi}{\lambda z} x^2 \right) \times \text{rect} \left(\frac{x}{2D} \right) \right\} \times \text{sinc}[(2p(f_x + f_c))] \times G(f_x) \} \quad (5.16)$$

The function $\text{sinc}(2pf_x)$ represents the pixel averaging effect in the Fourier domain with f_c as the carrier frequency. The bandwidth of object spectrum $\mathcal{F} \left\{ \exp \left(j \frac{\pi}{\lambda z} x^2 \right) \times \text{rect} \left(\frac{x}{2D} \right) \right\}$ is $\frac{2D}{\lambda z}$. This object spectrum is shifted f_c away from the zero frequency by off-axis geometry and becomes the 1st order spectrum in Fig. 5.3. In the reconstruction process, the product of $\text{sinc}(2pf_x)$ and the 1st order spectrum is shifted a distance $(-f_c)$ back to the centre in decoding process. This shift provides the term $\mathcal{F} \left\{ \exp \left(j \frac{\pi}{\lambda z} x^2 \right) \times \text{rect} \left(\frac{x}{2D} \right) \right\} \times \text{sinc}[(2p(f_x + f_c))]$ in Eq. (5.16). Finally the reconstructed image is obtained by performing a numerical back propagation from hologram/CCD plane to the object plane. This is achieved by multiplying the above term by $G(f_x)$ and then performing an inverse Fourier transform on the product. The process of Eq. (5.16) is simulated and described in detail in Fig. 5.4. The real part of normalized PSF is in the shape of $\text{sinc}(\frac{2D}{\lambda z} x)$. The imaginary part of normalized PSF is relatively small compared to the real part. Therefore the amplitude of normalized PSF is approximately equals to $\text{sinc}(\frac{2D}{\lambda z} x)$.

 Figure 5.3 Relation of $\text{sinc}(2pf_x)$ and 1st order in Fourier domain in the CCD or hologram plane.

Figure 5.4 Simulation of process of Eq. (5.16). Red indicates signal in Fourier domain and blue represents signal in spatial domain.

Hologram Stitching

In the hologram stitching process, three holograms at positions Δs , 0 , $-\Delta s$ are recorded by translating the compact digital holographic (CDH) system in one direction. The CCD aperture size increases from $2D$ to $2D+2\Delta s$ by stitching. This changes the PSF from $\text{sinc}(\frac{2D}{\lambda z}x)$ to $\text{sinc}(\frac{2D+2\Delta s}{\lambda z}x)$. The system simulation is done both before and after stitching to study its effect on the lateral resolution.

Lateral Resolution

Figure 5.5 (a) shows the PSFs for a CCD size $2D_x = 5.952\text{mm}$ in the x -direction, before stitching (solid) and after stitching (dotted) ($2D_x + 2\Delta s_x = 8.952\text{mm}$). Similarly for a CCD size $2D_y = 4.464\text{mm}$ in the y -direction, the PSF before stitching (solid) and after stitching (dotted) ($2D_y + 2\Delta s_y = 7.464\text{mm}$) are shown in Fig.5.5 (b). It can be seen that distance between the two first zeros of the sinc function decreases after stitching and therefore the lateral resolution of PSF is accordingly enhanced.

When we consider an object of finite extent L as shown in Fig. 5.6 (a), its spectrum at the CCD plane is the sum of the spectra of all the points along its extent according to the linearity theorem of Fourier transform and it extends from $-\frac{D}{\lambda z} - \frac{L}{2\lambda z}$ to $\frac{D}{\lambda z} + \frac{L}{2\lambda z}$ as in Fig. 5.6 (c) [146]. Hologram stitching expands the CCD size. Therefore the bandwidth $\frac{2D}{\lambda z}$ at each point in Fig. 5.6 (b) is expanded and the lateral resolution is improved. At the same time, the whole object spectrum extent $\frac{2D+L}{\lambda z}$ is also expanded. However, the highest spatial frequency which can be read by DH system is determined by $f_{max} = \frac{1}{2T} = \frac{1}{2\Delta x}$ where T is the sampling interval and Δx is the pixel size (The CCD fill factor is one here. The sampling interval is equal to pixel size). If object

spectrum exceeds this limit, the spectra overlap and boundary information of the object is lost. If CCD size expands Δs at each side, to avoid aliasing, the following condition should be satisfied:

$$\frac{2D+L+2\Delta s}{\lambda z} \leq \frac{1}{\Delta x} \quad (5.17)$$

Figure 5.5 Amplitude of normalized PSF amplitude before (solid) and after (dotted) stitching. (a) amplitude of normalized PSF with CCD size of $5.952mm$ (solid) and with CCD size $8.952mm$ (dotted) respectively; (b) amplitude of normalized PSF with CCD size of $4.464mm$ (solid) and with CCD size $7.464mm$ (dotted) respectively.

Figure 5.6 Spectrum of object with extent L_o at the CCD plane (before interference). (a) Object with extent L_o : B is at the center, A and C are the edges; (b) Spectrum of an arbitrary point located at x_i ; (c) Spectrum of object with extent L_o at the CCD plane.

5.3 Experiment and Results

5.3.1 Hologram Stitching

In the experiment, the CDH system (its photograph is shown in Fig. 5.7) is mounted on a motor controlled stage to record 9 holograms as in Fig. 5.8 with the main hologram located at the centre. Each hologram is $M \times N = 1280 \times 960$ pixels and each pixel is $\Delta x = 4.65 \mu\text{m}$ and $\Delta y = 4.65 \mu\text{m}$ in size. Holograms are shifted $\Delta s_x = 1.5 \text{ mm}$ and $\Delta s_y = 1.5 \text{ mm}$ away from the centre along the two axes. Before stitching, holograms are pre-processed as follows:

- In the current CDH system, the CCD, optics and the light source are integrated in one unit. When this unit is shifted to take holograms, the angle between reference beam and object beam may change by a small amount. Though small, this may cause a shift of the first order term in spectrum. A shift in the spectrum adds a linear phase tilt in wavefront. Therefore we cannot use a uniform reference wave to reconstruct all the holograms. In this experiment,

each 1st order spectrum is shifted to the centre individually to get the wavefront at the CCD to overcome the additional and undesired phase tilt. But in numerical reconstruction, the pixel at spectrum centre is of a width $\Delta f = \frac{1}{L}$ where L is the hologram size. Although all first orders are shifted to the spectrum centre, the maximum displacement error of object spectra is $\frac{1}{L}$ which corresponds to a 2π phase tilt in the whole field of view of the reconstructed phase image. To minimize this phase tilt, zero padding method can be used. According to our computer capability, each hologram is zero padded to 3 times of its original size. This reduces the tilt to less than $2\pi/3$ in the whole field of view.

- Even though phase tilts are compensated, there will be constant phase differences among the 9 phase images. To guarantee correct measurement, these phase differences should be suppressed. We use the sample substrate in the main hologram as a reference and let the substrates in the other areas share the same substrate phase/height as the main one.
- The calibration of each hologram position is quite important and necessary. Though the motor controlled moving stage is quite accurate, holograms positions still need to be calibrated. Displacement not only blurs the edges of intensity images, but may also causes destructive interference which makes the intensity of certain parts darker. The accuracy of phase information is also affected. Holograms are taken with overlapping in the recording process. We use the reconstructed intensity image of the main hologram at the centre as a reference and utilize the overlapping areas to calibrate the positions of other holograms. In this way, a position accuracy of $2.325\mu\text{m}$ in both x - direction and y -direction is achieved.

Figure 5.7 The photograph of CDH system.

Figure 5.8 Nine holograms taken by shifting CCD. The main hologram is in the middle.

Before stitching, The CCD is $2D_x = M \times \Delta x = 5.952mm$ in width and $2D_y = N \times \Delta y = 4.464mm$ in length. After stitching, CCD aperture is expanded to $2D_x + 2\Delta s_x = 8.952mm$ in width and $2D_y + 2\Delta s_y = 7.464mm$ in length. If Eq. (5.17) is satisfied, the lateral resolution of the system can be improved from $\frac{\lambda z}{Mag \times 2D_x}$ to $\frac{\lambda z}{Mag \times (2D_x + 2\Delta s_x)}$ in x -direction and from $\frac{\lambda z}{Mag \times 2D_y}$ to $\frac{\lambda z}{Mag \times (2D_y + 2\Delta s_y)}$ in y -direction according to Eq. (5.13).

5.3.2 Lateral Resolution and FOV Improvements

System parameters in this experiment are listed as below:

- $\lambda = 633nm$; $z = 119.7mm$; $Mag \cong 1.950$

- $\frac{\lambda z}{Mag \times 2D_x} = 6.528\mu\text{m}; \frac{\lambda z}{Mag \times 2D_y} = 8.704\mu\text{m}$
- $\frac{\lambda z}{Mag \times (2D_x + 2\Delta s_x)} = 4.341\mu\text{m}; \frac{\lambda z}{Mag \times (2D_y + 2\Delta s_y)} = 5.205\mu\text{m}$

We use an USAF (US Air Force) target as a lateral resolution target. The reconstructed intensity images before and after stitching are shown in Fig. 5.9. Figures (a), (b) and (c) are before stitching. (d), (e) and (f) are after stitching. (b) and (e) are the images in the highlighted square area of (a) and (d) respectively which present the G4 and G5 groups of the USAF target. (c) and (f) are the images in the highlighted square area of (b) and (e) respectively which show the G6 and G7 groups of the USAF target.

Figure 5.9 Reconstructed intensity images before and after stitching. (a), (b) and (c) images reconstructed from the main hologram (a single hologram); (d), (e) and (f) images reconstructed from the stitched hologram.

The lateral resolution is obviously enhanced as evident from the comparison of Fig. 5.9 (b) and (e), Fig. 5.9 (c) and (f). The FOV is simultaneously enlarged as seen by comparison of Fig. 5.9 (a) and (d). From a single hologram, the G6E2 lines can be resolved in x -direction and G5E6 lines can be resolved in y -direction which indicates that the practical lateral resolution of a single hologram is about $6.960\ \mu\text{m}$ and $8.769\ \mu\text{m}$ in x -direction and y -direction respectively. This agrees with the theoretical lateral resolution $\frac{\lambda z}{Mag \times 2D_x} = 6.528\ \mu\text{m}$ and $\frac{\lambda z}{Mag \times 2D_y} = 8.704\ \mu\text{m}$ in x -direction and y -direction very well.

After stitching, the G6E6 set in x -direction and the G6E5 set in y -direction are resolved. This indicates a lateral resolution of $4.385\ \mu\text{m}$ in x -direction and $4.922\ \mu\text{m}$ in y -direction. This agrees with the theoretical lateral resolution expectation -- $\frac{\lambda z}{Mag \times (2D_x + 2\Delta s_x)} = 4.341\ \mu\text{m}$ and $\frac{\lambda z}{Mag \times (2D_y + 2\Delta s_y)} = 5.205\ \mu\text{m}$ in x -direction and y -direction.

The above experiment demonstrates that hologram stitching has the capability to enhance the lateral resolution of DH system by increasing system aperture size. Furthermore it possesses the capability to simultaneously improve the lateral resolution and enlarge FOV.

5.4 Conclusion

In this chapter, it is shown for the first time that hologram stitching can simultaneously enhance the lateral resolution and enlarge the field of view (FOV) in an off-axis lensless Fresnel holography geometry. The impact of aperture synthesis of Fresnel holography on the lateral resolution is investigated with theoretical analysis and experiments. It is shown that as long as spectrum overlap is avoided, the lateral resolution can be improved at the same time as the

aperture extent is enlarged. The lensless Fresnel holographic geometry used in this work has a larger NA, and therefore better lateral resolution and compressed object spectrum compared to Fresnel holographic geometry with planar reference wave and object wave of the same reconstruction distance.

Compared with other aperture synthesis approaches [82, 103-134, 136, 137] used to improve the lateral resolution, this work has some strengths and weaknesses.

The first strength of this approach is the simultaneous improvement of lateral resolution and FOV. This strength comes from the working principle of this hologram stitching method. Each hologram contains both the low spatial frequencies diffracted by the object region of interest (ROI) and the high spatial frequencies diffracted by the areas around this ROI. In single hologram recording and reconstruction, only the ROI is reconstructed. This is due to the absence of low spatial frequencies of the surrounding area. Therefore, its intensity is much lower than the intensity of ROI and can barely be discerned. Hologram stitching enlarges the ROI such that the high spatial frequencies content of the region surrounding the original ROI can now be utilized for enhancing the lateral resolution. Hence, hologram stitching simultaneously improves both lateral resolution and FOV. In aperture synthesis using multiple illumination directions or specimen rotation, only the lateral resolution is enhanced [82, 112-137] without improvement of the FOV.

The second strength is the simple recording process particular for the compact digital holographic (CDH) system. The CDH can be easily mounted on a translation stage to record multiple holograms while maintaining the same incident angle of reference wave for all recordings. This simplifies the aperture synthesis process. For example, it is not necessary to

know the exact positions of holograms which can be automatically found by digital processing. However, in aperture synthesis approaches using multiple illumination directions or specimen rotation, the illumination directions and the object rotation angles need to be known in order to shift the high spatial frequencies to their original positions.

The third strength is the Fresnel holography configuration. Many works on aperture synthesis adopted Fourier holography configuration for simpler implementation [104, 106, 108, 112, 113, 116, 117, 138]. However Fourier configuration can only reconstruct image at one plane. Fresnel holography allows the reconstruction of all planes between object and CCD.

The weakness of this approach is that the maximum lateral resolution achieved is limited by the sampling interval of the CCD. The largest diffraction angle collected by CCD at each recording is determined by $\frac{1}{T}$ as discussed in the section of lateral resolution analysis. Since the reference and the illumination waves of this approach are kept constant in the recording of holograms at different positions, the upper limit of the lateral resolution of the stitched hologram is also determined by $\frac{1}{T}$. In the aperture synthesis approaches by multiple illumination directions and object rotation, the largest diffraction angle collected by CCD at each recording is determined by $\frac{1}{T}$ too. However, the collected spatial frequencies by changing the illumination directions or rotating the object will be moved back to their original high frequency locations according to the angle of illumination or the direction of object rotation in the synthesis process. Hence the upper limit of the lateral resolution is not limited by $\frac{1}{T}$ after aperture synthesis and is possible to reach the diffraction limit of the system as discussed in references [130, 131]. As reported in reference [130], there is a limit to the illumination direction -- grazing incidence. In order to approach the

lateral resolution defined by the diffraction limit, the approach of multiple illumination directions and the approach of object rotation are combined [130].

Table 5.1 concludes the strength and weakness of this work compared with other methods.

Strength	Weakness
Simultaneous improvement of lateral resolution and FOV	The maximum lateral resolution is limited by the spatial sampling interval of CCD
Simple recording process	
Fresnel holography configuration providing the capability for multi-plane reconstructions	

Table 5.1 The strengths and weakness of this work.

CHAPTER 6 ANALYSIS OF AXIAL MEASUREMENT ERRORS

6.1 Introduction

Resolution is a key parameter in evaluating the performance of a digital holographic system. The analysis and enhancement of lateral resolution have been thoroughly discussed in the Chapter 4 and Chapter 5. As a 3D metrology tool, the axial resolution is also very important for the system performance.

In digital holography, the axial dimension is measured by the Optical Path Difference (OPD) which is determined from the phase of the measured complex wavefront as $OPD = \lambda \frac{Phase}{2\pi}$. The relationship between the height of object and the OPD depends on the DH mode. In the reflection mode, the measured height is related to OPD as $Height = \frac{OPD}{2}$. For a transmission DH, the measured height is related to OPD as $Height = \frac{OPD}{n-n_0}$ where n is the refractive index of the object and n_0 is the refractive index of the surrounding medium. Therefore the OPD resolution is analogous to the axial resolution of the DH system. The corresponding axial resolution of different DH modes can be readily deduced from the OPD.

Although the CCD quantization effect is considered to provide an axial resolution limit, such resolution is hard to achieve in practice as reviewed in Table 3.3. One reason is the axial measurement errors may be much larger than the resolution set by quantization. There have been few researches on the factors which affect axial measurement, let alone thorough investigation

of the interaction between them and the weightages of the different factors on the measurement errors.

In a practical DH system, there are many potential parameters which may cause errors between the original object wavefront and the reconstructed object wavefront. The CCD camera used in the digital recording process introduces finite CCD size, pixel averaging due to the integration of the signal within a single pixel size and sampling effect. DH uses a reference wave for wavefront encoding and the conjugate of the reference wave for wavefront decoding. In the physical wave propagation and numerical reconstruction, wavelength and propagation distance which is also related to reconstruction distance also need to be considered. In numerical reconstruction, digital computing and reconstruction algorithms may introduce errors also. All these parameters will possibly influence the OPD/axial resolution.

In this work, the processes related to the CCD recording are discussed. Factors considered are the finite CCD size, pixel averaging and sampling effect due to CCD camera. A reference wave to produce a hologram at the CCD plane and the conjugate reference wave used to recover the wavefront from hologram are also discussed. Since the DH system is space variant, object position also affects the system performance. The impacts of the above factors on the OPD measurement accuracy are analyzed in this work.

This chapter is organized as follows. In section 6.2 the point spread function (PSF) with different limitations is derived and utilized to investigate the effects of these limitations in DH system. In section 6.3 the PSF for an ideal case is derived to provide a clearer view on the influences of these limitations. In section 6.4 the sampling effect is discussed. In section 6.5 a simulation investigation is performed. A step surface is used to investigate the influences of

finite CCD size, object displacement, the direction of reference wave and pixel averaging effect on the OPD resolution. In section 6.6 the influences of CCD size and object displacement on the axial accuracy is demonstrated with experiments. In section 6.7 the conclusions are given.

6.2 PSF of DH System

In the Fresnel diffraction integral, by including the finite CCD size, pixel averaging, sampling effects, reference wave and its conjugate and space variant property of practical DH, the reconstructed wavefront $Rf(x)$ of object wave field $f(x)$ can be written as:

$$Rf(x_2) = \text{Fresnel} \left[\left(\left[f(x) \otimes \exp\left(\frac{j\pi}{\lambda z} x^2\right) \right] \times \exp(-j2\pi ax) \times \text{rect}\left(\frac{x}{2D}\right) \right) \otimes \text{rect}\left(\frac{x}{2p}\right) \right] \times \exp(j2\pi ax) \times \sum_{-\infty}^{+\infty} \delta(x - nT) \quad (6.1)$$

where $\otimes$ denotes convolution, λ is wavelength used in recording, z is distance between object plane and CCD plane, $2D$ is CCD size, $2p$ is the pixel size and T is the sampling interval. The ratio $\frac{2p}{T}$ is the so-called fill factor. Function $\exp(-j2\pi ax)$ is the reference wave used in the hologram recording process. The corresponding tilt angle of the reference wave is $\theta = \sin^{-1}(a\lambda)$. In case a is zero, Eq. (6.1) is the reconstructed image of an in-line DH geometry. If a is not zero, Eq. (6.1) presents an off-axis geometry. We use x_2 to present the coordinate in image plane and x to present the coordinate on the CCD plane. *Fresnel* denotes the Fresnel transform given as:

$$\text{Fresnel}[k(x)] = \exp\left(-\frac{j\pi}{\lambda z} x_2^2\right) \mathcal{F}\{k(x) \times \exp\left(-\frac{j\pi}{\lambda z} x^2\right)\}_{f=-\frac{x_2}{\lambda z}} \quad (6.2)$$

Without loss of generality, a one-dimensional analysis is performed that can be scaled to two dimensions readily.

In Eq. (6.1), the convolution $f(x) \otimes \exp\left(\frac{j\pi}{\lambda z} x^2\right)$ describes the physical process of object wave propagation from object plane to CCD plane. At the CCD plane, the reference wave $\exp(-j2\pi ax)$ interferes with the object wave. This interference introduces the carrier frequency ‘ a ’ which shifts the carrier frequency of the object wavefront from zero to higher frequency position ‘ a ’ and ‘ $-a$ ’ in the Fourier domain. Only one of these object spectra is used in reconstruction which corresponds to the object wavefront with a tilt $\left[f(x) \otimes \exp\left(\frac{j\pi}{\lambda z} x^2\right)\right] \times \exp(-j2\pi ax)$ in spatial co-ordinates. The finite size of the CCD is the product of this tilted object wave with a rectangle function, $\text{rect}\left(\frac{x}{2D}\right)$. A further convolution with $\text{rect}\left(\frac{x}{2p}\right)$ introduces the pixel averaging effect over the whole CCD chip. The CCD sampling effect is a multiplication of wavefront with sampling signal $\sum_{-\infty}^{+\infty} \delta(x - nT)$.

Consider a point source $f(x) = \delta(x - x_0)$ at an arbitrary position x_0 in object plane, the PSF in the image plane is deduced from Eq. (6.1) and Eq. (6.2) as

$$PSF(x_2) = PSF_{x_0}(x_2) \otimes \sum_{n=-\infty}^{\infty} \delta(x_2 - \frac{n\lambda z}{T}) \quad (6.3)$$

with

$$PSF_{x_0}(x_2) =$$

$$\exp\left(-\frac{j\pi}{\lambda z} x_2^2\right) \times \mathcal{F} \left[\left(\exp\left[\frac{j\pi}{\lambda z} (x - x_0)^2\right] \times \text{rect}\left(\frac{x}{2D}\right) \times \exp(-j2\pi ax) \right) \otimes \text{rect}\left(\frac{x}{2p}\right) \times \exp(j2\pi ax) \times \exp\left(-\frac{j\pi}{\lambda z} x^2\right) \right]_{f=-\frac{x_2}{\lambda z}} \quad (6.4)$$

where $\mathcal{F}$ denotes Fourier transform. Using the commutative property of multiplication, operation sequence of $\text{rect}\left(\frac{x}{2D}\right)$ and $\exp(-j2\pi ax)$ is interchanged to facilitate analysis. The sampling effect in Eq. (6.3) generates multiple replicas of PSF_{x_0} with a $\frac{\lambda z}{T}$ interval. PSF_{x_0} is one of the

replica with $n=0$. In this section, interaction of finite CCD size, pixel averaging, reference wave tilt angle and point source position on PSF_{x_0} is discussed. The impact of sampling effect on their interactions is explored later in section 6.4. The various stages of Eq. (6.4) are shown graphically in Fig. 6.1 which helps to analyze the impact of these limitations. The left column is the function in the spatial domain (in blue) while the right column shows the function in the Fourier domain (in red).

Figure 6.1 Details of steps of Eq. (6.4).

Figure 6.2 The wavefront transformation sub-system at the hologram plane in DH system.

In the DH system, shown in Fig.6.2, at the hologram plane, the optical wavefront is transformed to digital wavefront in digital holography. This sub-system is called the wavefront transformation system. The input to this wavefront transformation system is the optical wavefront $\exp[\frac{j\pi}{\lambda z}(x - x_0)^2]$, which is the wavefront from source point $\delta(x_2 - x_0)$ impinging on the CCD. $F7$ in Fig. 6.1 is the output of this system. In an ideal system, $F7$ should be exactly equal to $\exp[\frac{j\pi}{\lambda z}(x - x_0)^2]$ which is not possible due to the constraints mentioned above in a practical system. The expression of $F7$ is quite complicated so that the errors are not straight forward. We will analyze the process from $\exp[\frac{j\pi}{\lambda z}(x - x_0)^2]$ to $F7$ of Eq. (6.4) with the help of Fig.6.1 to have a clear understanding of the relation between the errors and the constraints in both Fourier domain and spatial domain.

In the Fourier domain, the bandwidth of $\exp[\frac{j\pi}{\lambda z}(x - x_0)^2]$ is infinitely large. $F1$ in Fig. 6.1 is the object wavefront truncated by the finite CCD size, $2D$. Its spectrum $\mathcal{F}\{F1\}$ is shown in Fig. 6.1 (b) with bandwidth of $\frac{2D}{\lambda z}$. Within this bandwidth the spectrum amplitude is almost flat and is

equal to one and zero outside. It can be seen that the function of finite CCD size is to limit the bandwidth of object wavefront at the CCD from infinity to a finite size of $\frac{2D}{\lambda z}$. The introduction of reference wave $F2: \exp(-j2\pi ax)$ in the digital recording process is followed by the multiplication of the hologram with the conjugate of the recording reference wave $F6: \exp(j2\pi ax)$ in numerical reconstruction. From the amplitude spectra in Fig. 6.1 (d), it is observed that the reference wave $\exp(-j2\pi ax)$ shifts the object spectrum $\mathcal{F}\{F1\}$ by ‘ $-a$ ’ in frequency domain. The pixel averaging effect is the convolution of $F4: \text{rect}\left(\frac{x}{2p}\right)$ in space which results in a $\text{sinc}(2pf)$ modulation of the shifted object spectrum as in Fig. 6.1 (h). After this, the conjugate of the reference wave $F6: \exp(j2\pi ax)$ shifts the object spectrum back by ‘ a ’ in spectrum in Fig. 6.1 (j). It can be clearly seen in Fig. 6.1 (j), the intertwined interaction of the above three factors is a modulation of function $\text{sinc}\{2p(f - a)\}$ on the object spectrum $\mathcal{F}\{F1\}$ and arrives at wavefront F7 in space. The spectrum in F7 is $\text{sinc}\{2p(f - a)\} \times \mathcal{F}\{F1\}$ function.

From the above analysis, the errors between the input wavefront $\exp\left[\frac{j\pi}{\lambda z}(x - x_0)^2\right]$ and output digital wavefront F7 in the Fourier domain include a bandwidth limitation of $\frac{2D}{\lambda z}$ due to the finite CCD size and a modulation $\text{sinc}\{2p(f - a)\}$ on the bandwidth limited object spectrum due to the interaction of reference wave, pixel averaging effect and the conjugate of the reference wave.

As F7 is the wavefront of an arbitrary point located at x_o , the influence of object position is also included in it. Numerical reconstruction is then performed on digital wavefront F7 by a Fresnel transform. This Fresnel transform numerically propagates the wavefront F7 at the CCD plane to the image plane and the reconstructed image PSF is obtained as seen in Fig. 6. 1 (k) to (n).

In the spatial domain, the complex amplitude of F7 can be expressed as

$$F7 = \exp\left[\frac{j\pi}{\lambda z}(x - x_0)^2\right] \times \text{sinc}\left[\frac{2p}{\lambda z}(x - x_0 - \lambda za)\right] \times \text{rect}\left(\frac{x}{2D}\right) \quad (6.5)$$

The errors in the optical wavefront and digitized wavefront in space domain are seen in the comparison of $\exp\left[\frac{j\pi}{\lambda z}(x - x_0)^2\right]$ and Eq. (6.5). Finite CCD size introduces modulation of $\text{rect}\left(\frac{x}{2D}\right)$ to the wavefront $\exp\left[\frac{j\pi}{\lambda z}(x - x_0)^2\right]$. The pixel averaging effect, the reference wave and its conjugate add a $\text{sinc}\left[\frac{2p}{\lambda z}(x - x_0 - \lambda za)\right]$ term to the input wavefront.

The errors of wavefront $F7$ at the hologram plane result in measurement errors at the image. If the wavefront $\exp\left[\frac{j\pi}{\lambda z}(x - x_0)^2\right]$ at the hologram plane is propagated to the image plane, we get $\delta(x_2 - x_0)$ as an image. However if $F7$ at the hologram plane is propagated to the image plane, all the constraints on it result in errors between the reconstructed image and $\delta(x_2 - x_0)$ at the image plane. The influences of these constraints in the image, especially on phase, are investigated here. By substituting Eq. (6.5) into Eq. (6.4) we get

$$\begin{aligned} PSF_{appr, x_0}(x_2) &= \exp\left(-\frac{j\pi}{\lambda z}x_2^2\right) \times \mathcal{F}\text{resnel}\{F7\}_{f=-\frac{x_2}{\lambda z}} \\ &= \exp\left(-\frac{j\pi}{\lambda z}x_2^2\right) \times \mathcal{F}\left\{\exp\left[\frac{j\pi}{\lambda z}(x - x_0)^2\right] \times \text{sinc}\left[\frac{2p}{\lambda z}(x - x_0 - \lambda za)\right] \times \text{rect}\left(\frac{x}{2D}\right) \times \exp\left(-\frac{j\pi}{\lambda z}x^2\right)\right\}_{f=-\frac{x_2}{\lambda z}} \\ &= \exp\left[-\frac{j\pi}{\lambda z}(x_2^2 - x_0^2)\right] \times \mathcal{F}\left\{\text{sinc}\left[\frac{2p}{\lambda z}(x - x_0 - \lambda za)\right] \times \text{rect}\left(\frac{x}{2D}\right) \times \exp\left[-j\frac{2\pi}{\lambda z}xx_0\right]\right\}_{f=-\frac{x_2}{\lambda z}} \end{aligned} \quad (6.6)$$

$$= \exp\left[-\frac{j\pi}{\lambda z}(x_2^2 - x_0^2)\right] \times \left\{ \left[\exp\left(-j2\pi\frac{x_0 + \lambda za}{2p}x_2\right) \times \text{rect}\left(\frac{x}{2p}\right) \right] \otimes \text{sinc}\left(\frac{2D}{\lambda z}x_2\right) \otimes \delta(x_2 - x_0) \right\} \quad (6.7)$$

The phase of PSF_{appr, x_0} has two parts: one is factor $\exp\left[-\frac{j\pi}{\lambda z}(x_2^2 - x_0^2)\right]$ and the other is term $\mathcal{F}\left\{\text{sinc}\left[\frac{2p}{\lambda z}(x - x_0 - \lambda za)\right] \times \text{rect}\left(\frac{x}{2D}\right) \times \exp\left[-j\frac{2\pi}{\lambda z}xx_0\right]\right\}_{f=-\frac{x_2}{\lambda z}}$ in Eq. (6.6). In the first part, the phase is related to the coordinate x_2 in the reconstruction plane and the source point position x_0 . In the second part, the factor $\exp\left[-j\frac{2\pi}{\lambda z}xx_0\right]$ does not contribute to the phase in the

image since it corresponds to a shift of the image. The remaining $\text{sinc}[\frac{2p}{\lambda z}(x - x_o - \lambda za)] \times \text{rect}(\frac{x}{2D})$ contributes to errors in images.

Here we mainly investigate the influences on phase, as phase errors cause axial measurement errors. Since the phase is determined by the imaginary and real part of a complex wave, we discuss how the imaginary part and real part relate to factor $\mathcal{F}\{\text{sinc}[\frac{2p}{\lambda z}(x - x_o - \lambda za)] \times \text{rect}(\frac{x}{2D})\}_{f=-\frac{x_2}{\lambda z}}$. Any real function can be expressed as the sum of a real even function and a real odd function. Real function $\text{sinc}[\frac{2p}{\lambda z}(x - x_o - \lambda za)] \times \text{rect}(\frac{x}{2D})$ can be expressed by the sum of a real even function $g(x)$ and a real odd function $h(x)$ as

$$\text{sinc}[\frac{2p}{\lambda z}(x - x_o - \lambda za)] \times \text{rect}(\frac{x}{2D}) = g(x) + h(x) \quad (6.8)$$

with

$$g(x) = \frac{1}{2} \left\{ \text{sinc}[\frac{2p}{\lambda z}(x - x_o - \lambda za)] \times \text{rect}(\frac{x}{2D}) + \text{sinc}[\frac{2p}{\lambda z}(x + x_o + \lambda za)] \times \text{rect}(\frac{x}{2D}) \right\} \quad (6.9)$$

$$h(x) = \frac{1}{2} \left\{ \text{sinc}[\frac{2p}{\lambda z}(x - x_o - \lambda za)] \times \text{rect}(\frac{x}{2D}) - \text{sinc}[\frac{2p}{\lambda z}(x + x_o + \lambda za)] \times \text{rect}(\frac{x}{2D}) \right\} \quad (6.10)$$

According to the properties of Fourier transform, the Fourier transform of real even function and real odd function are real even function and pure imaginary odd function respectively. At the position $x_2 = x_o$ where the source point locates, the phase comes from factor $\exp\left[-\frac{j\pi}{\lambda z}(x_2^2 - x_o^2)\right]$ is zero. Since $\mathcal{F}\{h(x)\}$ is an odd function, the phase of factor $\mathcal{F}\{\text{sinc}[\frac{2p}{\lambda z}(x - x_o - \lambda za)] \times \text{rect}(\frac{x}{2D}) \times \exp[-j\frac{2\pi}{\lambda z}xx_o]\}_{f=-\frac{x_2}{\lambda z}}$ at $x_2 = x_o$ is zero too. Thus, at $x_2 = x_o$, PSF in Eq. (6.6) is real positive with zero phase. Away from this position, the phase of the PSF is not zero and the value changes with position x_2 and is space variant. When comparing the PSF in Eq. (6.6) with $\delta(x_2 - x_o)$ which is the PSF of ideal system, $\delta(x_2 - x_o)$ has a positive real value and

zero phase at $x_2 = x_0$. But away from $x_2 = x_0$, the phase of $\delta(x_2 - x_0)$ is also zero. Thus the axial measurement errors root from the non-zero imaginary part of the normalized PSF induced by the constraints above. To investigate the influences of different factors on the imaginary part of PSF in Eq. (6.6) and (6.7), the imaginary part in the main lobe of the normalized PSF is considered. In Eq. (6.7), the amplitude of normalized PSF is a *sinc* function centered at $x_2 = x_0$. The influence of each term on the average imaginary value (AIV) and the maximum imaginary value (MIV) in the main lobe is studied by varying the term of interest while maintaining the other terms at the same value.

- **Influence of Finite CCD size**

The influence of finite CCD size on the imaginary part of the PSF is investigated by increasing the CCD size from 2.3808mm to 16.666mm in steps of 2.3808mm. Fig. 6.3 illustrates the dependence of AIV and MIV on the CCD size. The pixel size is $2p = T = 4.65\mu m$. The carrier frequency of reference wave is $a = \frac{1}{4T} = 53.76 \times 10^3 Hz$ and the corresponding tilt angle is $\theta = \sin^{-1}(a\lambda) = 1.95^\circ$. The point source is located at the center of the object plane.

Figure 6.3 The relationship between AIV, MIV and CCD size 2D.

It is seen that the AIV and MIV increase with larger CCD size. The reason for this is that larger CCD aperture increases the asymmetry and therefore increases the proportion of $h(x)$ in Eq. (6.8). The larger the proportion of the component $h(x)$ is in the sum of Eq. (6.8), the greater the imaginary part in $\mathcal{F}\{\text{sinc}[\frac{2p}{\lambda z}(x - x_o - \lambda za)] \times \text{rect}(\frac{x}{2D})\}_{f=-\frac{x_2}{\lambda z}}$. As the term $\exp\left[-\frac{j\pi}{\lambda z}(x_2^2 - x_0^2)\right]$ is unchanged, the increase of the imaginary part $\mathcal{F}\{\text{sinc}[\frac{2p}{\lambda z}(x - x_o - \lambda za)] \times \text{rect}(\frac{x}{2D})\}_{f=-\frac{x_2}{\lambda z}}$ contributes to the increase in the imaginary part of the PSF.

• Influence of Point Source Position

The influence of point source position on the PSF imaginary part is investigated by translating the object point from the center of the object plane (0 mm) to a distance of 1.7856 mm in steps of 0.2976 mm. Fig. 6.4 illustrates the relationship between AIV and MIV and the position. The

CCD size is $2D = 4.7616\text{mm}$. The pixel size is $2p = T = 4.65\mu\text{m}$ and the carrier frequency is

$a = \frac{1}{4T} = 53.76 \times 10^3\text{Hz}$ which corresponds to a tilt angle $\theta = \sin^{-1}(a\lambda) = 1.95^\circ$.

Normalized imaginary part

Figure 6.4 The relationship between AIV, MIV and point source position x_0 .

It is seen that the AIV and MIV increase as the point source position x_0 is further away from the center of object plane. In this case, both phase factor $\exp\left[-\frac{j\pi}{\lambda z}(x_2^2 - x_0^2)\right]$ and $\mathcal{F}\{\text{sinc}[\frac{2p}{\lambda z}(x - x_0 - \lambda za)] \times \text{rect}\left(\frac{x}{2D}\right)\}_{f=-\frac{x_2}{\lambda z}}$ change with position. In the first term greater the x_0 larger the oscillation of the real and imaginary parts. In this simulation, the position of the point source x_0 and the carrier frequency a are both positive. Increasing of x_0 increases the asymmetry of $\text{sinc}[\frac{2p}{\lambda z}(x - x_0 - \lambda za)] \times \text{rect}\left(\frac{x}{2D}\right)$ and therefore increases the imaginary part of $\mathcal{F}\{\text{sinc}[\frac{2p}{\lambda z}(x - x_0 - \lambda za)] \times \text{rect}\left(\frac{x}{2D}\right)\}_{f=-\frac{x_2}{\lambda z}}$. Both of the two phase terms increase the imaginary part of the PSF as the point source moves away from the origin.

- **Influence of Carrier Frequency (Reference wave direction)**

The influence of the carrier frequency a introduced by the angle of the reference wave on the PSF imaginary part is investigated by changing the carrier frequency from 0 to $\frac{1}{2T}$ in steps of $\frac{1}{16T}$ where $T = 4.65\mu m$. The angle correspondingly changes from 0 to 3.9° in steps of 0.487° . Fig. 6.5 illustrates the relationship between the AIV and MIV and the carrier frequency. The CCD size is $2D = 4.7616mm$, the pixel size is $2p = T = 4.65\mu m$. The point source is located at the center of the object plane.

Figure 6.5 The relationship between AIV, MIV and carrier frequency of reference wave a .

It is observed that the AIV and MIV increase with increase in the carrier frequency. The factor $\exp\left[-\frac{j\pi}{\lambda z}(x_2^2 - x_0^2)\right]$ in PSF is unchanged in this case. The object spectrum $\mathcal{F}\{F1\}$ is an even function and symmetrical with respect to y -axis. In the case carrier frequency is not zero, the object spectrum limited by pixel averaging: $\text{sinc}[2p(f - a)] \times \mathcal{F}\{F1\}$ is not symmetrical anymore. A higher carrier frequency enlarges the asymmetry of the imaginary part of

$\text{sinc}[\frac{2p}{\lambda z}(x - x_o - \lambda za)] \times \text{rect}(\frac{x}{2D})$ and therefore increases the imaginary part of $\mathcal{F}\{\text{sinc}[\frac{2p}{\lambda z}(x - x_o - \lambda za)] \times \text{rect}(\frac{x}{2D})\}_{f=-\frac{x_2}{\lambda z}}$ and PSF.

• Influence of Pixel Averaging Effect

The influence of the pixel averaging on the imaginary part of PSF is investigated for a pixel size $2p$ that varies from $\frac{1}{8}T$ to T in steps of $\frac{1}{8}T$ where $T = 4.65\mu m$. The corresponding fill factor $\frac{2p}{T}$ changes from $\frac{1}{8}$ to 1 in steps of $\frac{1}{8}$. Fig. 6.6 illustrates the relationship between the AIV and MIV and the fill factor. The CCD size is $2D = 4.7616mm$. The carrier frequency is $a = \frac{1}{4T} = 53.76 \times 10^3 Hz$ and the corresponding tilt angle is $\theta = \sin^{-1}(a\lambda) = 1.95^\circ$. The point source is located at the center of the object plane.

Figure 6.6 The relationship between AIV, MIV and fill factor $\frac{2p}{T}$.

It is seen that the larger pixel size or fill factor the higher the AIV and MIV. The factor $\exp\left[-\frac{j\pi}{\lambda z}(x_2^2 - x_0^2)\right]$ in PSF is unchanged in this case. The object spectrum $\mathcal{F}\{F1\}$ is an even function and symmetrical with respect to the y -axis. Increasing the pixel size enlarges the asymmetry character of $\text{sinc}\left[\frac{2p}{\lambda z}(x - x_0 - \lambda za)\right] \times \text{rect}\left(\frac{x}{2D}\right)$ and therefore increases the imaginary part of $\mathcal{F}\left\{\text{sinc}\left[\frac{2p}{\lambda z}(x - x_0 - \lambda za)\right] \times \text{rect}\left(\frac{x}{2D}\right)\right\}_{f=-\frac{x_2}{\lambda z}}$ and hence increases the imaginary part of PSF.

From the above analysis, it can be seen that a larger CCD, larger distance of object point from center, higher carrier frequency and greater fill factors all increase the imaginary part of PSF.

6.3 Ideal Cases

In above section 6.2, the PSF with all the limitations is derived and investigated. Idealized situations such as infinitely large CCD size and infinitely small pixel size are derived for a clearer understanding in this section. This will assist in developing schemes for accuracy enhancement.

6.3.1 Infinitely large CCD Size and Infinitely Small Pixel Size

If the CCD size is infinitely large and the pixel size is infinitely small, then $\text{rect}\left(\frac{x}{2D}\right) \rightarrow 1$ and $\text{rect}\left(\frac{x}{2p}\right) \rightarrow \delta(x)$. The PSF_{x_0} of DH system in Eq. (6.4) in such case becomes:

$$\begin{aligned} PSF_{x_0, 2p \rightarrow 0, 2D \rightarrow +\infty}(x_2) &= \exp\left(-\frac{j\pi}{\lambda z}x_2^2\right) \times \mathcal{F}\left[\exp\left[\frac{j\pi}{\lambda z}(x - x_0)^2\right] \times \exp\left(-\frac{j\pi}{\lambda z}x^2\right)\right]_{f=-\frac{x_2}{\lambda z}} \\ &= \exp\left[-\frac{j\pi}{\lambda z}(x_2^2 - x_0^2)\right] \times \delta(x_2 - x_0) = \delta(x_2 - x_0) \end{aligned} \quad (6.11)$$

which is exactly same as the point source object. In this case, the DH system is a space invariant and both the lateral resolution and the phase/OPD resolution are infinitely small.

6.3.2 Infinitely Small Pixel Size

For an infinitely small pixel, we have $rect\left(\frac{x}{2p}\right) \rightarrow \delta(x)$. The PSF_{x_0} of DH system in Eq. (6.4)

in this case is

$$\begin{aligned} PSF_{x_0, 2p \rightarrow 0}(x_2) &= \exp\left(-\frac{j\pi}{\lambda z} x_2^2\right) \times \mathcal{F}\left[\exp\left[\frac{j\pi}{\lambda z}(x - x_0)^2\right] \times rect\left(\frac{x}{2D}\right) \times \exp\left(-\frac{j\pi}{\lambda z} x^2\right)\right]_{f=-\frac{x_2}{\lambda z}} \\ &= \exp\left[-\frac{j\pi}{\lambda z}(x_2^2 - x_0^2)\right] \times \left[sinc\left(\frac{2D}{\lambda z} x_2\right) \otimes \delta(x_2 - x_0)\right] \\ &= \exp\left[-\frac{j\pi}{\lambda z}(x_2^2 - x_0^2)\right] \times sinc\left[\frac{2D}{\lambda z}(x_2 - x_0)\right] \end{aligned} \quad (6.12)$$

Despite the phase term $\exp\left[-\frac{j\pi}{\lambda z}(x_2^2 - x_0^2)\right]$, $sinc\left[\frac{2D}{\lambda z}(x_2 - x_0)\right]$ is a real function. The phase in the main lobe of $sinc\left[\frac{2D}{\lambda z}(x_2 - x_0)\right]$ is zero. Hence the phase in the main lobe of PSF_{x_0} is determined by $\exp\left[-\frac{j\pi}{\lambda z}(x_2^2 - x_0^2)\right]$. In this case, the DH system is a space variant system. The lateral resolution is determined by $\frac{\lambda z}{2D}$. At $x_2 = x_0$, the phase is zero. But for $x_2 \neq x_0$ within the main lobe, the phase is $-\frac{\pi}{\lambda z}(x_2^2 - x_0^2)$.

6.3.3 Infinitely Large CCD Size

For an infinitely large CCD, $rect\left(\frac{x}{2D}\right) \rightarrow 1$. The PSF_{x_0} in Eq. (6.4) can be written as:

$$\begin{aligned} PSF_{2D \rightarrow +\infty, x_0}(x_2) &= \exp\left(-\frac{j\pi}{\lambda z} x_2^2\right) \times \\ &\mathcal{F}\left[\left\{\exp\left[\frac{j\pi}{\lambda z}(x - x_0)^2\right] \times \exp(-j2\pi ax)\right\} \otimes rect\left(\frac{x}{2p}\right)\right] \times \exp(j2\pi ax) \times \exp\left(-\frac{j\pi}{\lambda z} x^2\right) \Big|_{f=-\frac{x_2}{\lambda z}} \end{aligned} \quad (6.13)$$

To explain the interactions of the pixel averaging, reference wave and point source position on PSF_{x_0} , we investigate Eq. (6.13) in relation with Fig. 6.7. Processes from $F2$ to $F6$ are not described here but with the conclusion obtained in section 6.2 that the intertwined interaction of reference wave, pixel averaging effect and the conjugate of reference wave is a modulation of function $\text{sinc}\{2p(f - a)\}$ on the object spectrum $\mathcal{F}\{F1\}$. $F7$ is the object wavefront with the interaction of the pixel averaging effect and the reference wave and its conjugate of DH system. As $F7$ is the wavefront of an arbitrary point located at x_o , the influence of object position is also included in it. Due to the infinitely large CCD, the spectrum $\mathcal{F}\{F1\}$ is infinitely large. Therefore, in this case, the object bandwidth is determined by $\frac{1}{p}$ which stems from modulation function $\text{sinc}\{2p(f - a)\}$. The complex amplitude of wavefront $F7$ of Eq. (6.5) can be expressed as:

$$F7_{2D \rightarrow +\infty} = \exp\left[\frac{j\pi}{\lambda z}(x - x_o)^2\right] \times \text{sinc}\left[\frac{2p}{\lambda z}(x - x_o - \lambda za)\right] \quad (6.14)$$

The $PSF_{x_0, 2D \rightarrow +\infty}$ in Eq. (6.13) can be expressed approximately by $PSF_{appr, x_0, 2D \rightarrow +\infty}$

$$\begin{aligned} PSF_{appr, x_0, 2D \rightarrow +\infty}(x_2) &= \exp\left(-\frac{j\pi}{\lambda z}x_2^2\right) \times \mathcal{F}\text{resnel}\{F7_{2D \rightarrow +\infty}\}_{f=-\frac{x_2}{\lambda z}} \\ &= \exp\left[-\frac{j\pi}{\lambda z}(x_2^2 - x_o^2)\right] \times \mathcal{F}\{\text{sinc}\left[\frac{2p}{\lambda z}(x - x_o - \lambda za)\right] \times \exp[-j\frac{2\pi}{\lambda z}xx_o]\}_{f=-\frac{x_2}{\lambda z}} \\ &= \exp\left[-\frac{j\pi}{\lambda z}(x_2^2 - x_o^2)\right] \times \left\{\left[\exp\left(-j2\pi\frac{x_o + \lambda za}{2p}x_2\right) \times \text{rect}\left(\frac{x_2}{2p}\right)\right] \otimes \delta(x_2 - x_o)\right\} \\ &= \exp\left[-\frac{j\pi}{\lambda z}(x_2^2 - x_o^2)\right] \times \left\{\exp\left[-j2\pi\frac{x_o + \lambda za}{2p}(x_2 - x_o)\right] \times \text{rect}\left(\frac{x_2 - x_o}{2p}\right)\right\} \end{aligned} \quad (6.15)$$

which is a complex function. The lateral resolution is determined by $2p$. The phase factor $\exp\left[-\frac{j\pi}{\lambda z}(x_2^2 - x_o^2)\right] \times \exp\left[-j2\pi\frac{x_o + \lambda za}{2p}(x_2 - x_o)\right]$ is valid only in the range $\left|\frac{x_2 - x_o}{2p}\right| \leq \frac{1}{2}$ due to the multiplication with $\text{rect}\left(\frac{x_2 - x_o}{2p}\right)$. Therefore, the PSF_{x_0} of DH system with infinitely large CCD size has an extent in space of size $2p$. Within this extent the phase is not zero and equal to

$-\pi \left[\frac{(x_2^2 - x_0^2)}{\lambda z} + \frac{x_0 + \lambda z a}{p} (x_2 - x_0) \right]$ which is space variant. The phase error and hence the axial measurement error are determined by the combination of factors including x_2 , x_0 , a , λ , z and $2p$.

Figure 6.7 Investigation of steps of Eq. (6.13).

6.3.4 Summary

The PSF_{x_0} under the influences of different constraints is investigated in section 6.2 and 6.3. Based on this investigation of PSF, we discuss the influence of constraints on an object $f(x)$ with certain extent L . This object $f(x)$ can be considered as an integration of all the points on it and expressed as

$$f(x) = \int_{-\frac{L}{2}}^{\frac{L}{2}} A(x) \delta(x - x_0) dx_0 = \int_{-\frac{L}{2}}^{\frac{L}{2}} A(x_0) \delta(x - x_0) dx_0 \quad (6.16)$$

where $A(x)$ is the amplitude of $f(x)$. With Eq. (6.16), properties of convolution and the linearity of Fourier transform, a replica of the reconstructed image of Eq. (6.1) can be expanded as

$$R_0 f(x_2) = \exp\left(-\frac{j\pi}{\lambda z} x_2^2\right) \times \int_{-\frac{L}{2}}^{\frac{L}{2}} A(x_0) \times \mathcal{F}\left[\left\{\exp\left[\frac{j\pi}{\lambda z}(x - x_0)^2\right] \times \text{rect}\left(\frac{x}{2D}\right) \times \exp(-j2\pi ax)\right\} \otimes \text{rect}\left(\frac{x}{2p}\right)\right] \times \exp(j2\pi ax) \times \exp\left(-\frac{j\pi}{\lambda z} x^2\right) \Bigg|_{f=-\frac{x_2}{\lambda z}} dx_0 \quad (6.17)$$

In the case with all constraints, the reconstructed image $R_0 f(x_2)$ can be expressed by the PSF in Eq. (6.7) and shown in Table 6.1 case 1. When the pixel size becomes infinitely small, the reconstructed image $R_0 f_{2p \rightarrow 0}(x_2)$ can be derived as shown in Table 6.1 case 2 by setting $\text{rect}\left(\frac{x}{2p}\right) \rightarrow \delta(x)$ in Eq. (6.17). When the CCD size becomes infinitely large, the reconstructed image $R_0 f_{2D \rightarrow +\infty}(x_2)$ is acquired as shown in Table 1 case 3 by setting $\text{rect}\left(\frac{x}{2D}\right) \rightarrow 1$ in Eq. (6.17). When the CCD size is infinitely large and the pixel size is infinitely small, the reconstructed image $R_0 f_{2p \rightarrow 0, 2D \rightarrow +\infty}(x_2)$ is shown in Table 1 case 4 obtained by setting $\text{rect}\left(\frac{x}{2p}\right) \rightarrow \delta(x)$ and $\text{rect}\left(\frac{x}{2D}\right) \rightarrow 1$ in Eq. (6.17).

Case 1:	$R_0 f(x_2) = \int_{-\frac{L}{2}}^{\frac{L}{2}} A(x_0) \times \exp \left[-\frac{j\pi}{\lambda z} (x_2^2 - x_0^2) \right] \times \left\{ \exp \left[-j2\pi \frac{x_0 + \lambda z a}{2p} (x_2 - x_0) \right] \times \text{rect} \left(\frac{x_2 - x_0}{2p} \right) \right\} \otimes \text{sinc} \left[\frac{2D}{\lambda z} (x_2 - x_0) \right] dx_0$
Case 2: $2p \rightarrow 0$	$R_0 f_{2p \rightarrow 0}(x_2) = \int_{-\frac{L}{2}}^{\frac{L}{2}} A(x_0) \times \exp \left[-\frac{j\pi}{\lambda z} (x_2^2 - x_0^2) \right] \times \text{sinc} \left[\frac{2D}{\lambda z} (x_2 - x_0) \right] dx_0$
Case 3: $2D \rightarrow +\infty$	$R_0 f_{2D \rightarrow +\infty}(x_2) = \int_{-\frac{L}{2}}^{\frac{L}{2}} A(x_0) \times \exp \left[-\frac{j\pi}{\lambda z} (x_2^2 - x_0^2) \right] \times \exp \left[-j2\pi \frac{x_0 + \lambda z a}{2p} (x_2 - x_0) \right] \times \text{rect} \left(\frac{x_2 - x_0}{2p} \right) dx_0$
Case 4: $2p \rightarrow 0,$ $2D \rightarrow +\infty$	$R_0 f_{2p \rightarrow 0, 2D \rightarrow +\infty}(x_2) = \int_{-\frac{L}{2}}^{\frac{L}{2}} A(x_0) \times \exp \left[-\frac{j\pi}{\lambda z} (x_2^2 - x_0^2) \right] \times \delta(x_2 - x_0) dx_0 = \int_{-\frac{L}{2}}^{\frac{L}{2}} A(x_0) \delta(x_2 - x_0) dx_0$

Table 6.1 The reconstructed images under the influences of different constraints.

Comparing these results, it can be seen that a system with an infinitely large CCD size and infinitely small pixel size can fully reconstruct the object as the original image without any loss of information (Table 6.1 case 4). But for finite CCD and pixel size, in addition to image smoothing, there are phase errors. A finite CCD size introduces a convolution of $\text{sinc} \left(\frac{2D}{\lambda z} x_2 \right)$ with the image. For any arbitrary point $A(x_0) \delta(x_2 - x_0)$, the PSF of this system is $A(x_0) \times \exp \left[-\frac{j\pi}{\lambda z} (x_2^2 - x_0^2) \right] \times \text{sinc} \left[\frac{2D}{\lambda z} (x_2 - x_0) \right]$. Away from $x = x_0$ the phase error is spatially dependent. Pixel size results in a convolution of $\exp \left(-j2\pi \frac{x_0 + \lambda z a}{2p} x_2 \right) \times \text{rect} \left(\frac{x_2}{2p} \right)$ with the image. For any arbitrary point $A(x_0) \delta(x_2 - x_0)$, the PSF is $A(x_0) \times \exp \left[-\frac{j\pi}{\lambda z} (x_2^2 - x_0^2) \right] \times \exp \left[-j2\pi \frac{x_0 + \lambda z a}{2p} (x_2 - x_0) \right] \times \text{rect} \left(\frac{x_2 - x_0}{2p} \right)$. Away from x_0 and in the range $|x_2 - x_0| \leq p$, the phase is related to the interaction of x_0 , $2p$ and a . For any arbitrary point $A(x_0) \delta(x_2 - x_0)$ of system with both finite CCD size and pixel size, the PSF of this system becomes $A(x_0) \times \exp \left[-\frac{j\pi}{\lambda z} (x_2^2 - x_0^2) \right] \times \left\{ \left[\exp \left(-j2\pi \frac{x_0 + \lambda z a}{2p} x_2 \right) \times \text{rect} \left(\frac{x_2}{2p} \right) \right] \otimes \text{sinc} \left(\frac{2D}{\lambda z} x_2 \right) \otimes \delta(x_2 - x_0) \right\}$ in which both the factor $\text{sinc} \left(\frac{2D}{\lambda z} x_2 \right)$ introduced by the CCD size $2D$ and the factor $\exp \left(-j2\pi \frac{x_0 + \lambda z a}{2p} x_2 \right) \times \text{rect} \left(\frac{x_2}{2p} \right)$ caused by the pixel averaging effect, the carrier frequency a and spatial variant factor x_0 convolute with the image. How these factors affect the axial

measurement error is more complex to see from the convolution. We investigate their interactions on axial error in section 6.5.

6.4 Sampling Effect

In order to avoid spectrum aliasing, the following condition has to be satisfied:

$$\frac{2D}{\lambda z} + \frac{L_0}{\lambda z} \leq \frac{1}{T} \quad (6.18)$$

6.5 Investigation of Axial Measurement Errors of Object with Extent

In this section, we simulate the influence of object position, pixel size, CCD size and carrier frequency on axial measurement errors. Since axial measurement is derived from the OPD as discussed in section 6.1, the OPD error is used as a measure of axial measurement error.

There are several reasons for a simulation based analysis. Firstly in practical system, some factors are not easy to change over the desired range. Secondly changes in the parameter of interest may affect other parameters due to the limitation of system setup. Thirdly besides parameters of interest, other factors such as noise may affect the performance.

The object used in simulation has a step profile represent by a solid red line in Fig. 6.8. In the investigation of each parameter, the reconstructed images of this object are calculated using Eq. (6.17) by only changing the parameter of interest and determining the OPD error.

The step height of the simulated object is 50 nm resulting in an OPD of 100 nm between the step surface and the substrate. The substrate is 297.6 μm wide on either side of the step which is 595.2 μm in width. The wavelength used is $\lambda = 633 \text{ nm}$ and the reconstruction distance z is 139.91 mm. The sampling interval T is 4.65 μm .

Figure 6.8 Step object (red solid line) and its reconstructed image (blue dotted line).

The reconstructed image of the object is shown by the blue dotted line in Fig. 6.8. The OPD error of the reconstructed image from the original object can be clearly seen with greatest error at the sharp edges of the object. Three parameters are defined to present the characters of the OPD errors. The two black dotted lines in Fig. 6.8 identify the maximum OPD error position. The average of the two maximum OPD error values is used to indicate the maximum OPD error (MOE) of the step surface. The averaged OPD errors between these two black dotted lines are used to indicate the average OPD error (AOE) of the step surface. The green dotted line lies at the maximum error at the step edge on the substrate side. The distance between the maximum OPD error point on the step and the maximum OPD error point on the substrate of the same edge is defined as the edge width (EW). With parameters MOE, AOE and EW, the influences of finite CCD size, object position, carrier frequency and pixel size on the OPD measurement are investigated.

6.5.1 Influence of Finite CCD Size

The influence of finite CCD size on the OPD measurement error is investigated for a CCD whose size increases from 2.3808mm to 16.666mm in steps of 2.3808mm. The pixel size is $2p = T = 4.65\mu m$. The carrier frequency is $a = \frac{1}{4T} = 53.76 \times 10^3 Hz$ and the corresponding angle between the object and reference beams is $\theta = \sin^{-1}(a\lambda) = 1.95^\circ$. The object is located at the centre of the CCD. Fig. 6.9 illustrates the dependence of MOE, AOE and EW on the change of CCD size. It is seen that a larger CCD helps to reduce the AOE, MOE and EW and thus improve the OPD measurement accuracy.

The reason for EW reduction is that the lateral resolution of DH system is determined by term $\text{sinc}\left(\frac{2D}{\lambda z}x_2\right)$. To avoid aliasing, object spectrum $\frac{2D+L_0}{\lambda z}$ should be smaller than $\frac{1}{T}$. Also since T is greater than $2p$, the relation $\frac{2D}{\lambda z} < \frac{1}{2p}$ is valid. In such case, the factor CCD size dominates over the pixel size in determining the lateral resolution as discussed in chapter 4. Hence a larger CCD provides better lateral resolution. The relation between lateral resolution and AOE, MOE and EW is shown in Fig. 6.10. The EW is nearly equal to the lateral resolution, in Fig. 6.10 (c).

CHAPTER 6 ANALYSIS OF AXIAL MEASUREMENT ERRORS

Figure 6.9 The relationship between AOE, MOE and EW and CCD size.

Figure 6.10 The relationship between (a) AOE; (b) MOE; (c) EW and lateral resolution $\frac{\lambda z}{D}$.

It can also be seen that a larger CCD with better resolution translates to smaller OPD error in Fig. 6.9 and Fig. 6.10. This is because larger CCD can collect diffracted signal of larger angle which corresponds to higher frequencies of the object structure. In such case less information of higher orders are lost and hence the axial error of image becomes smaller.

6.5.2 Influence of Object Position

The influence of object position on the OPD measurement is investigated by translating the object from -1.488mm to 1.488mm at object plane in steps of 0.2976mm. Parameters that remain unchanged are $2D = 4.7616\text{mm}$, $2p = T = 4.65\mu\text{m}$ and $a = \frac{1}{4T} = 53.76 \times 10^3\text{Hz}$.

Figure 6.11 (a) The object (red solid) and its reconstructed OPD image (blue dotted) in case that the object is located symmetrically at the center of object plane; the object (red solid) and its reconstructed OPD image (blue dotted) in case the object is located -0.4464 mm away from center of object plane; (c) Relationship between DMSE and the object position.

It is noticed that the once the object position is moved away from the center, the OPD error become unsymmetrical as seen in the comparison of Fig. 6.11 (a) and (b). In (a) the object (red solid line) with reconstructed OPD (blue dotted line) is located symmetrically with respect to the optical axis at the object plane. In (b) the object (red solid line) with its reconstructed OPD image (blue dotted line) is located asymmetrically with respect to the optical axis at the object plane. To identify the extent of asymmetry of the OPD image, we define the difference of the mean square error of the left side and the right side of the step surface (DMSE) as a factor to investigate the relationship between the asymmetry and the object position. From Fig. 6.11 (c), it can be seen that larger the object displacement larger the asymmetry of the reconstructed OPD image.

Considering the width of $\text{sinc}\left(\frac{2D}{\lambda z}x_2\right)$ in Eq. (6.7) with a main lobe width $\frac{\lambda z}{D}$, the PSF is valid over the range:

$$|x_2 - x_0| \leq x_0 + \frac{\lambda z}{D} + p \quad (6.19)$$

When $x_2 = x_0$, the phase errors from the two terms $\exp\left[-j\frac{\pi}{\lambda z}(x_2^2 - x_0^2)\right] = \exp\left[-j2\pi\frac{x_2+x_0}{2\lambda z}(x_2 - x_0)\right]$ and $\exp\left[-j2\pi\frac{x_0+\lambda za}{2p}(x_2 - x_0)\right]$ are both zero. The phase error of PSF happens at $x \neq x_0$ in the valid range of Eq. (6.19). The former phase factor plays a major role in the phase error. As $\frac{1}{2\lambda z}$ is about 10^3 times of $\frac{1}{2p}$ in practice, the slope of the first factor $2\pi\frac{x_2+x_0}{2\lambda z}$ is more sensitive than that of the second factor $2\pi\frac{x_0+\lambda za}{2p}$. Furthermore the slope $2\pi\frac{x_2+x_0}{2\lambda z}$ is related with x_0 . When x_0 is larger, the phase error of PSF is larger. When the object is not symmetrically located, points located further away from the centre have larger OPD error. Hence in the reconstructed image which is the sum of all the PSF on the object, the OPD error is asymmetrical and longer side has larger OPD error. Alternately, the

asymmetry in OPD error of off-axis object is due to the space variant property of digital holography.

6.5.3 Influence of Carrier Frequency

The influence of the carrier frequency α results from an angle between the object wave and reference wave. The OPD measurement is studied for carrier frequency ranging from 0 to $\frac{1}{2T}$ in steps of $\frac{1}{16T}$. The corresponding angle ranges from 0 to 3.9° in steps of 0.487° . Fig. 6.12 illustrates the relationship between MOE, AOE and EW and the change in carrier frequency. The parameters which do not change are $2D = 4.7616mm$ and $2p = T = 4.65\mu m$. The object is located at the center.

It is observed from Fig. 6.12 that higher the carrier frequency larger the AOE and MOE of the measurement. EW is not affected by the carrier frequency. This indicates that higher carrier frequency results in larger OPD error and hence reduces the OPD measurement accuracy. But the slop of the step edge is not dependent on the carrier frequency.

Figure 6.12 Relationship between (a) AOE; (b) MOE; (c) EW and the carrier frequency a .

We analyze this result with the help of spectral domain. To avoid overlapping, the bandwidth of object spectrum $\mathcal{F}\{F1\}$ should be smaller than $\frac{1}{T}$. The bandwidth of $\text{sinc}\{2p(f - a)\}$ is larger than $\frac{2}{T}$ as fill factor $\frac{2p}{T}$ cannot exceed 1. Furthermore a has to be smaller than $\frac{1}{2T}$ to avoid aliasing. Therefore the object spectrum $\mathcal{F}\{F1\}$ lies between the two first zeros of $\text{sinc}\{2p(f - a)\}$.

The reason for the increase of AOE and MOE is that larger carrier frequency a enlarges the asymmetry extent of $\text{sinc}[2p(f - a)] \times \mathcal{F}\{F1\}$. Once the object is symmetrically placed, $\mathcal{F}\{F1\}$ is a symmetrical even function. When the carrier frequency is not zero, the object spectrum $\text{sinc}[2p(f - a)] \times \mathcal{F}\{F1\}$ is not symmetrical. Higher carrier frequency leads to larger asymmetry of $\text{sinc}[2p(f - a)] \times \mathcal{F}\{F1\}$ resulting in a larger ratio of imaginary part to the real part of image that increases the phase and OPD measurement error. As the change of carrier

frequency only shifts the modulation $\text{sinc}[2p(f - a)]$ on the object spectrum $\mathcal{F}\{F1\}$, the bandwidth of object spectrum is not affected. Therefore the lateral resolution and hence EW is not affected.

6.5.4 Influence of Pixel Size

The influence of the pixel size on the OPD measurement error is investigated by varying the pixel size from $\frac{1}{8}T$ to T in steps of $\frac{1}{8T}$ where $T = 4.65\mu\text{m}$. The corresponding fill factor $\frac{2p}{T}$ changes from $\frac{1}{8}$ to 1 in steps of $\frac{1}{8}$. Fig. 6.13 illustrates the variations in MOE, AOE and EW for different fill factors. The parameters that do not change are $2D = 4.7616\text{mm}$ and $a = \frac{1}{4T} = 53.76 \times 10^3\text{Hz}$. The object is located at the center.

Figure 6.13 relationship between (a) AOE; (b) MOE; (c) EW and the fill factor $\frac{2p}{T}$.

It is seen that the larger fill factor reduces the AOE, MOE of the measurement while EW is not affected. This suggests that larger fill factor reduces OPD measurement error but the slope of the step is not influenced by fill factor.

As discussed in Table 6.1 case 1, the pixel size intertwines with carrier frequency and point position in the convolution $\left\{ \exp \left[-j2\pi \frac{x_0 + \lambda za}{2p} (x_2 - x_0) \right] \times \text{rect} \left(\frac{x_2 - x_0}{2p} \right) \right\} \otimes \text{sinc} \left[\frac{2D}{\lambda z} (x_2 - x_0) \right]$. Convolution results in a smoothing effect. A larger fill factor not only provides a larger smoothing effect but also reduce the rate of phase change $\frac{1}{2p}$ in factor $\exp \left[-j2\pi \frac{x_0 + \lambda za}{2p} (x_2 - x_0) \right]$. Both of them help to reduce the OPD error.

In the spectral domain, the modulation of $\text{sinc}\{2p(f - a)\}$ on the object spectrum $\mathcal{F}\{F1\}$ does not affect the bandwidth of object spectrum. Hence, the lateral resolution and the EW is not obviously affected by the change of pixel size. However this modulation $\text{sinc}\{2p(f - a)\}$ on the object spectrum $\mathcal{F}\{F1\}$ changes the energy distribution. The energy of higher frequencies is suppressed and more energy is concentrated toward lower frequencies with larger $2p$. As a result, more energy is included in the main lobe of the PSF in space and the side lobes of PSF are suppressed to lower amplitude. Hence the PSF of one point is less affected by the side lobes of the PSFs of nearby points. The accumulated OPD error due to the tails of other point is less. Therefore the reconstructed image which is the weighted sum of all PSF along the object has lower OPD error. This modulation process is similar to the concept of Gaussian filtering.

6.6 Experiment and Results

In this part the influences of finite CCD size and object displacement on the axial measurement errors are experimentally examined. There are experimental difficulties when investigating the influences of the carrier frequency and pixel size. Experimentally, once the reference beam angle is changed, other parameters are also affected due to system geometry. Furthermore, since the influence of the carrier frequency is small, its influence may be masked by the influence of other changes. Therefore the result obtained may not be convincing. The pixel size of the CCD cannot be changed and hence the effect of pixel size on axial measurement is not performed experimentally.

6.6.1 Finite CCD Size

In this experiment, seven holograms are recorded with different CCD sensing sizes, achieved by blocking a part of the CCD during recording. The sizes of the seven holograms are 960×960 , 910×910 , 860×860 , 810×810 , 760×760 , 710×710 , 660×660 pixels respectively. Each pixel is of $4.65 \mu\text{m} \times 4.65 \mu\text{m}$ in size for all cases.

An USAF (US Air Force) target as shown in Fig. 6.14 is the object. USAF target is a standard resolution chart and the heights of all the coated bars (the bright parts in Fig. 6.14) are 100nm. This is a good and readily available sample to study the axial measurement accuracy.

Figure 6.14 USAF target.

Figure 6.15 The reconstructed height images from holograms with size (a) 960×960 ; (b) 910×910 ; (c) 860×860 ; (d) 810×810 ; (e) 760×760 ; (f) 710×710 ; (g) 660×660 pixels respectively.

The reconstructed height images from holograms of different sizes are shown in Fig. 6.15. As the effective field of view is different in each reconstruction, to investigate the axial measurement with different holograms, a common area -- group 2 element 3 (G2E3) as highlighted by the black squares of images is used for further investigation.

CHAPTER 6 ANALYSIS OF AXIAL MEASUREMENT ERRORS

The 3D profile images of G2E3 element reconstructed for different hologram sizes are shown in Fig. 6.16.

Figure 6.16 The reconstructed 3D height images of G2E3 from holograms with size (a) 960×960; (b) 910×910; (c) 860×860; (d) 810×810; (e) 760×760; (f) 710×710; (g) 660×660 pixels respectively.

a. Investigation of Average Axial Error

To investigate the axial measurement error, the average height error (AHE) of the G2E3 element is evaluated. The AHE of G2E3 element in Fig. 6.16 is calculated. The AHE values in the seven reconstructions are shown in Fig. 6.17. It can be seen that a larger CCD size reduces the AHE and therefore improves the axial accuracy. This result agrees with the investigation of AOE and CCD in Fig. 6.9 (a) obtained in the simulation section 6.5.1.

The absolute values of AHE in Fig. 6.17 are larger than the absolute values of AOE in Fig. 6.9. One reason is that, besides the factors discussed in the simulation part, there are other factors contributing to the axial errors in experiments. Furthermore, the simulation of Fig. 6.9 is performed in one dimension. The error will become larger if two dimensions are considered. This can also explain why the variation range of AHE is also larger than that of AOE in Fig. 6.9.

Figure 6.17 The relationship between AHE value and CCD size.

b. Investigation of Maximum Axial Error

CHAPTER 6 ANALYSIS OF AXIAL MEASUREMENT ERRORS

The maximum height error (MHE) of G2E3 element is also evaluated for the axial accuracy analysis.

As seen in Fig. 6.13, the object in experiment has two dimensions rather than the one dimension in the simulation. Each bar is considered as an integration of multiple section planes. For each section plane, there is a maximum error at each edge as illustrated in an example shown in Fig. 6.18. These two maximum errors are recorded at each plane. The average value of the maximum errors of all the section planes is calculated as the MHE. The MHE values for different CCD size are calculated and plotted in Fig. 6.19. The relation between the MHE and CCD size is presented. It can be seen that larger CCD size can help to reduce MHE and hence can improve axial measurement accuracy. This result agrees with the result obtained in the simulation investigation of section 6.5.1.

Figure 6.18 (a) The 3D height map of the a bar; (b) height profile at section plane 1 in (a); (c) height profile at section plane 2 in (a).

Figure 6.19 The relationship between MHE value and CCD size.

c. Investigation of Edge Width

Edge width (EW) of the bar is used to investigate the measurement ability of steep jump in axial direction (z direction). As can be seen in Fig. 6.16, each bar has four edges. EW is the average edge width of all edges of G2E3 element. EW values are calculated for the seven different CCD sizes.

The relation between the EW and CCD size is shown in Fig. 6.20. The solid line is EW values measured from the experiment. The dashed lines shows the lateral resolution range predicted by the system parameters at different hologram sizes. From Fig. 6.20, it can be seen that EW follows the trend of theoretical lateral resolution of DH system. This demonstrates that larger CCD size can help to reduce EW and hence can improve the measurement of axial steep jump. This result agrees with the result obtained in the simulation investigation of EW in section 6.5.1 where the EW value is nearly equal to the lateral resolution $R = \frac{\lambda z}{2D}$. The EW measured by experiment (solid line in Fig. 6.20), though larger than $R = \frac{\lambda z}{2D}$, is in the range from R to $1.1xR$ (dashed lines). This agrees with the conclusion of chapter 4 that the lateral resolution should be in the range from $\frac{\lambda z}{2D}$ to $1.1x\frac{\lambda z}{2D}$ with the according system parameters in an off-axis geometry.

Figure 6.20 Relation between the EW and CCD size. The yellow solid line is EW measured with experiment. The green dashed line is the theoretical resolution at different hologram sizes.

As a summary of the experiment on the influence of CCD size. It can be clearly seen that the results acquired from experiments agree with the conclusions obtained in simulation part of section 6.5.1. Larger CCD size can improve axial measurement accuracy and the ability for the sharp edge width measurement.

6.6.2 Object Displacement

As stated in the simulation part of section 6.5.2, object should be placed in the center of object plane so that the object is symmetrically located with respect to the optical axis. Displacement away from the center causes asymmetry in the axial measurement error.

In this part, the axial measurement errors are evaluated with different values of object displacement from center. Holograms with object displacement ranging from -1.53mm to 1.53mm in steps of 0.51mm are shown in Fig. 6.21 (G2E3 element is again selected). The

reconstructed images of holograms in Fig. 6.21 are shown in Fig. 6.22. The profiles of a bar are shown in Fig. 6.22 to see the change of asymmetry in the axial measurement at different displacement values. The changes at the edges (emphasized by circles in Fig. 6.22) are especially obvious.

Figure 6.21 Holograms with object displacement of (a) 1.53mm; (b) 1.02mm; (c) 0.51mm; (d) 0mm; (e) -0.51mm; (f) -1.02mm; (g) -1.53mm.

To quantitatively describe the asymmetry of the axial measurement error, the difference of the mean square error (DMSE) of the left side and the right side of the bar is calculated as in the simulation part. The relationship between DMSE and the object displacement is shown in Fig. 6.23. It can be seen that larger object displacement with respect to the optical axis results in larger asymmetry of the axial measurement error. This result agrees with the one obtained in simulation part of section 6.5.2. Therefore object should be placed in the center of the field of view for better measurement.

Figure 6.22 Reconstructed height images from holograms with object displacement of (a) 1.53mm; (b) 1.02mm; (c) 0.51mm; (d) 0mm; (e) -0.51mm; (f) -1.02mm; (g) -1.53mm. In each image, profile at the same section plane of the bar is shown to see the change of asymmetry in the axial measurement at different the displacement values. Circles are used to emphasize the change at the edge.

Figure 6.23 Relationship between the DMSE and the object displacement.

6.7 Conclusion

In this chapter, the impacts of system constraints related to practical digital holography on the axial measurement accuracy are investigated. These system constraints include the tilt angle of reference wave in the hologram recording, finite CCD size, the pixel averaging effect within pixel size and sampling effect introduced in AOC process of hologram. As DH system is space variant, the effect of object displacement is also included. PSF is utilized to analyze the influences on the system performance first. After that the system performance of axial measurement accuracy is investigated.

In the analysis of PSF, we find that the point source position, finite CCD size, pixel size and carrier frequency individually introduce factors $\delta(x_2 - x_0)$, $\text{sinc}\left(\frac{2D}{\lambda z}x_2\right)$ and $\exp\left(-j2\pi\frac{x_0+\lambda za}{2p}x_2\right) \times \text{rect}\left(\frac{x_2}{2p}\right)$ respectively. The combination of different constraints causes convolution between corresponding factors as in Eq. (6.11) to Eq. (6.15) and in table 6.1.

From the investigation of axial/OPD measurement accuracy, it can be concluded that larger CCD size can reduce the OPD measurement error and can make the measurement of sharp edge width more accurate. Larger pixel size and smaller carrier frequency help to reduce the OPD measurement error but does not affect the measurement of sharp edge width. The object displacement has an impact on the symmetry of the reconstructed image. When object is not symmetrically placed with respect to the optical axis, the asymmetry in the measured OPD error increases. Comparing the impact extent of all these factors, finite CCD size has the largest impact on the OPD error. Furthermore, it affects the sharp edge measurement accuracy. The second largest impact factor is the asymmetry of the object placement. Sampling effect causes a limitation of CCD size and object size in a digital holography as in Eq. (6.18). Therefore, though larger CCD size can improve OPD measurement performance, it cannot be enlarged as desired with object of fixed size due to the sampling effect.

The influences of CCD size and object displacement on the axial accuracy is further demonstrated with experiments. The results obtained from experiments agree with the conclusions achieved by the simulation investigation.

The conclusions achieved in this chapter can be used as a guide for DH system adjustment to improve the axial measurement accuracy.

CHAPTER 7 CONCLUSION AND FUTUER WORK

7.1 Conclusion

Digital holography is a quantitative 3D imaging and measurement technique. Practically its lateral resolution and axial measurement accuracy are limited by systematic constraints. In this thesis, the lateral resolution and its improvements are studied. The axial measurement errors are investigated.

The relationship between the phase of object wave and the object 3D profile in digital holography is introduced in the beginning of the thesis. This is necessary for investigation of the lateral resolution and axial accuracy that are discussed later.

A survey of current phase imaging and measurement techniques is performed. The working principles of different techniques are analyzed. Their advantages and shortcomings are discussed and compared. Digital holography shows better capabilities in aspects of direct and quantitative phase imaging, simple system configuration, fast, better axial resolution and multiple-plane reconstruction etc.

The general introduction of digital holography technique is provided. Different configurations are introduced. Digital recording and numerical reconstruction processes are explained in details. After numerical reconstruction, the digital object wavefield is obtained. This makes the numerical correction of aberrations possible. A review of the approaches for aberration compensation is provided. The lateral resolution and axial accuracy are very important parameters of digital holography. The research works of lateral resolution analysis, its

CHAPTER 7 CONCLUSIONS AND FUTUER WORK

improvement and the investigation of axial accuracy is reviewed. Based on the reviews, possible exploration directions are point out.

Compared to conventional holography, the lateral resolution of digital holography is limited due to the CCDs or other recording devices. The lateral resolution of digital holography limited by CCD introduced factors is analyzed. Three factors contribute to this limitation, namely, pixel averaging effect within the finite detection size of a single pixel, finite CCD aperture size and sampling effect due to finite sampling interval. As DH system is space variant, influences of object extent on lateral resolution are also involved. Interactions of these factors on lateral resolution are investigated and presented. The lateral resolution of DH system can be determined for given parameters of these factors. The domains dominated by different factors are explained along with their accuracy. Lateral resolution performance of in-line and off-axis systems is also studied and examples of lateral resolution determination for a practical system are provided.

With the results on the lateral resolution analysis, the improvement of lateral resolution is investigated by increasing the numerical aperture of the system with a different aperture synthesis method. Both the lateral resolution and image field of view can be enhanced at the same time using a more general Fresnel holography setup by hologram stitching. In the experiment, the synthesis is executed by moving the compact digital holographic system in two directions. Nine holograms are recorded and stitched into one hologram. The reconstruction results show that expanding aperture can improve lateral resolution.

As a 3D measurement technique, the axial measurement performance of digital holography is also very important. The axial measurement accuracy of digital holography under the influences of different constraints is analyzed. The systematic processes related to the hologram recording

are focused on. Constraints related to finite CCD size, pixel averaging due to the integral of signal within single pixel detection size and sampling effect due to CCD camera are examined. The influence of the reference wave used for object wave encoding into the hologram and its conjugate for object wave decoding from the hologram is also investigated. Object position also affects the system performance. The impact of all the above factors on the axial measurement errors are analyzed thoroughly in this part. This investigation can be used as a guide for system adjustment to improve the axial measurement accuracy.

The main contributions of this thesis are summarized as follow:

- A survey and a comparison of current phase measurement techniques
- A general introduction of digital holography
- Literature review on the studies of lateral resolution, its improvement and axial measurement accuracy
- Investigation of the limitation of system factors on the lateral resolution in digital holography:
 - Finite CCD aperture size
 - Averaging effect within the finite detection size of one pixel
 - Sampling effect
 - Space variant property of DH system
- Simultaneous improvement of lateral resolution and FOV by hologram stitching approach
- Analysis of the system limitations on the axial accuracy in digital holography:
 - Finite CCD aperture size
 - Object Position

- Reference Wave
- Averaging effect within the finite detection size of one pixel

7.2 Future Works

There is always room for further enhancement and improvements.

Though hologram stitching can simultaneously improve the lateral resolution and FOV, the maximum lateral resolution achieved is limited by the sampling interval of CCD. The other two approaches by multiple illumination directions and object rotation do not have this limitation. But they are not able to provide simultaneous improvement of the lateral resolution and FOV. Therefore these approaches can be combined to integrate their advantages and overcome shortcomings. By hologram stitching, larger FOV and lateral resolution approaching the limit set by the CCD sampling can be achieved. To go beyond this limit and approach the lateral resolution set by diffraction limit, illumination angles and the object orientation can be changed to record multiple holograms at each recording position. Since the spatial frequencies lower than the sampling limitation is obtained by hologram stitching, a large part of works which are previously performed by illumination angels and large object rotation are saved. By a suitable aperture synthesis method, simultaneous improvement of FOV and lateral resolution approaching the diffraction limit can be accomplished. Indeed, there is the further need to see if methods of super-resolution for sub-diffraction imaging can be achieved as well in a lens-less geometry.

CHAPTER 7 CONCLUSIONS AND FUTUER WORK

The influences of four system factors on the axial accuracy are investigated with both simulation and experiment. The finite CCD size and object displacement are found to have the largest impacts on the axial accuracy of digital holographic system among the four factors discussed. However besides these, there are other possible factors which influence the axial resolution too. These factors include object size, object height, reconstruction algorithms and the quantization effect of CCD, etc. A further investigation of these factors is meaningful.

In the study of chapter 6, it is found that with larger CCD size and placing the object symmetrically with respect to the optical axis can improve the axial accuracy. The nature of the axial error caused by these two factors is the ringing effect. Approaches to reduce or eliminate the ringing effect can help to reduce the axial error. The exploration of these approaches is necessary and meaningful.

REFERENCES

1. Nathan S. Claxton, T.J.F., and Michael W. Davidson *Laser Scanning Confocal Microscopy*
2. Wikipedia. *Confocal Microscopy*. Available from: http://en.wikipedia.org/wiki/Confocal_microscope.
3. *Website of Nikon's MicroscopyU*. Available from: <http://www.microscopyu.com>.
4. Barone-Nugent, E.D., A. Barty, and K.A. Nugent, *Quantitative phase-amplitude microscopy I: optical microscopy*. Journal of Microscopy-Oxford, 2002. **206**: p. 194-203.
5. Teague, M.R., *DETERMINISTIC PHASE RETRIEVAL - A GREEN-FUNCTION SOLUTION*. Journal of the Optical Society of America, 1983. **73**(11): p. 1434-1441.
6. Kazuo Ishizuka, B.A., *Defocus Step of Trough-Focus Images used to Phase Measurement in the Transport of Intensity Equation*, in *8th Asia-Pacific Conference on Electron Microscopy*. 2004: Kanazawa, Japan.
7. Paganin, D. and K.A. Nugent, *Noninterferometric phase imaging with partially coherent light*. Physical Review Letters, 1998. **80**(12): p. 2586-2589.
8. Allen, L.J. and M.P. Oxley, *Phase retrieval from series of images obtained by defocus variation*. Optics Communications, 2001. **199**(1-4): p. 65-75.
9. Barty, A., et al., *Quantitative optical phase microscopy*. Optics Letters, 1998. **23**(11): p. 817-819.
10. Curl, C.L., et al., *Quantitative phase microscopy: A new tool for investigating the structure and function of unstained live cells*. Clinical and Experimental Pharmacology and Physiology, 2004. **31**(12): p. 896-901.
11. Curl, C.L., et al., *Quantitative phase microscopy: a new tool for measurement of cell culture growth and confluency in situ*. Pflugers Archiv-European Journal of Physiology, 2004. **448**(4): p. 462-468.
12. Curl, C.L., et al., *Single cell volume measurement by quantitative phase microscopy (QPM): A case study of erythrocyte morphology*. Cellular Physiology and Biochemistry, 2006. **17**(5-6): p. 193-200.
13. Curl, C.L., et al., *Refractive index measurement in viable cells using quantitative phase-amplitude microscopy and confocal microscopy*. Cytometry Part A, 2005. **65A**(1): p. 88-92.
14. Ampem-Lassen, E., et al., *Refractive index profiling of axially symmetric optical fibers: a new technique*. Optics Express, 2005. **13**(9): p. 3277-3282.
15. Zernike, F., *Phase-contrast, a new method for microscopic observation of transparent objects. Part II*. Physica: 9, 1942: p. 974-986.
16. Zernike, F., *Phase-contrast, a new method for microscopic observation of transparent objects. Part I*. Physica: 9, 1942: p. 686-698.
17. Zernike, F., *Diffraction theory of the knife-edge test and its improved form, the phase-contrast method*. Royal Astronomy Society Monthly Notices: 94, 1934: p. 377-384.
18. Lang, W., *Nomarski differential interference contrast microscopy. I. Fundamentals and experimental designs*. Zeiss Information, 70, Carl Zeiss, Oberkochen, Germany, 1968: p. 114-120.
19. Lang, W., *Nomarski differential interference contrast microscopy. II. Formation of the interference image*. Zeiss Information, 71, Carl Zeiss, Oberkochen, Germany, 1969: p. 12-16.
20. Oldfield, R., *Light Microscopy, An Illustrated Guide*. 1994: Wolfe, London.
21. Wikipedia. *Differential interference contrast microscopy*. Available from: http://en.wikipedia.org/wiki/Differential_interference_contrast_microscopy.
22. Popescu, G., et al., *Fourier phase microscopy for investigation of biological structures and dynamics*. Optics Letters, 2004. **29**(21): p. 2503-2505.
23. Lue, N., et al. *Quantitative phase imaging of live cells using fast Fourier phase microscopy*. in *Biomedical Optics Topical Meeting of the Optical-Society-of-America*. 2006. Ft Lauderdale, FL: Optical Soc Amer.

REFERENCES

24. Ikeda, T., et al., *Hilbert phase microscopy for investigating fast dynamics in transparent systems*. Optics Letters, 2005. **30**(10): p. 1165-1167.
25. Popescu, G., et al., *Erythrocyte structure and dynamics quantified by Hilbert phase microscopy*. Journal of Biomedical Optics, 2005. **10**(6).
26. Lue, N., et al., *Tissue refractometry using Hilbert phase microscopy*. Optics Letters, 2007. **32**(24): p. 3522-3524.
27. Popescu, G., et al., *Diffraction phase microscopy for quantifying cell structure and dynamics*. Optics Letters, 2006. **31**(6): p. 775-777.
28. Park, Y., et al., *Fresnel particle tracing in three dimensions using diffraction phase microscopy*. Optics Letters, 2007. **32**(7): p. 811-813.
29. Popescu, G., et al. *Imaging red blood cell dynamics by quantitative phase microscopy*. in *Red Cell Club Meeting*. 2007. Boston, MA: Academic Press Inc Elsevier Science.
30. Park, Y.K., et al., *Diffraction phase and fluorescence microscopy*. Optics Express, 2006. **14**(18): p. 8263-8268.
31. Lue, N., et al., *Confocal diffraction phase microscopy of live cells*. Optics Letters, 2008. **33**(18): p. 2074-2076.
32. Cogswell, C.J., et al. *Quantitative DIC microscopy using a geometric phase shifter*. in *Conference on Three-Dimensional Microscopy - Image Acquisition and Processing IV*. 1997. San Jose, Ca.
33. Kou, S.S. and C.J.R. Sheppard. *Quantitative phase restoration in differential interference contrast (DIC) microscopy - art. no. 700005*. in *Conference on Optical and Digital Image Processing*. 2008. Strasbourg, FRANCE.
34. Kouskousis, B., et al., *Quantitative phase and refractive index analysis of optical fibers using differential interference contrast microscopy*. Applied Optics, 2008. **47**(28): p. 5182-5189.
35. Van Munster, E.B., L.J. Van Vliet, and J.A. Aten, *Reconstruction of optical pathlength distributions from images obtained by a wide-field differential interference contrast microscope*. Journal of Microscopy-Oxford, 1997. **188**: p. 149-157.
36. Choma, M.A., et al., *Spectral-domain phase microscopy*. Optics Letters, 2005. **30**(10): p. 1162-1164.
37. Joo, C., et al., *Spectral-domain optical coherence phase microscopy for quantitative phase-contrast imaging*. Optics Letters, 2005. **30**(16): p. 2131-2133.
38. Choma, M.A., et al., *Sensitivity advantage of swept source and Fourier domain optical coherence tomography*. Optics Express, 2003. **11**(18): p. 2183-2189.
39. Wojtkowski, M., et al., *In vivo human retinal imaging by Fourier domain optical coherence tomography*. Journal of Biomedical Optics, 2002. **7**(3): p. 457-463.
40. Marquet, P., et al., *Digital holographic microscopy: a noninvasive contrast imaging technique allowing quantitative visualization of living cells with subwavelength axial accuracy*. Optics Letters, 2005. **30**(5): p. 468-470.
41. Jueptner, U.S.a.W., *Digital Holography*. 2004.
42. Schnars, U., *DIRECT PHASE DETERMINATION IN HOLOGRAM INTERFEROMETRY WITH USE OF DIGITALLY RECORDED HOLOGRAMS*. Journal of the Optical Society of America a-Optics Image Science and Vision, 1994. **11**(7): p. 2011-2015.
43. Schnars, U. and W. Juptner, *DIRECT RECORDING OF HOLOGRAMS BY A CCD TARGET AND NUMERICAL RECONSTRUCTION*. Applied Optics, 1994. **33**(2): p. 179-181.
44. Schnars, U. and W.P.O. Juptner, *Digital recording and numerical reconstruction of holograms*. Measurement Science & Technology, 2002. **13**(9): p. R85-R101.
45. GOODMAN, J.W., *Introduction to Fourier Optics*. 2nd ed. 1996: McGRAW-HILL.
46. Matsushima, K., H. Schimmel, and F. Wyrowski, *Fast calculation method for optical diffraction on tilted planes by use of the angular spectrum of plane waves*. Journal of the Optical Society of America a-Optics Image Science and Vision, 2003. **20**(9): p. 1755-1762.
47. Grilli, S., et al., *Whole optical wavefields reconstruction by digital holography*. Optics Express, 2001. **9**(6): p. 294-302.

REFERENCES

48. Xu, L., J.M. Miao, and A. Asundi, *Properties of digital holography based on in-line configuration*. Optical Engineering, 2000. **39**(12): p. 3214-3219.
49. Yamaguchi, I. and T. Zhang, *Phase-shifting digital holography*. Optics Letters, 1997. **22**(16): p. 1268-1270.
50. Yamaguchi, I., et al., *Image formation in phase-shifting digital holography and applications to microscopy*. Applied Optics, 2001. **40**(34): p. 6177-6186.
51. Xu, L., et al., *Hybrid holographic microscope for interferometric measurement of microstructures*. Optical Engineering, 2001. **40**(11): p. 2533-2539.
52. Xu, L., et al., *Imaging analysis of digital holography*. Optics Express, 2005. **13**(7): p. 2444-2452.
53. Xu, L., et al., *Studies of digital microscopic holography with applications to microstructure testing*. Applied Optics, 2001. **40**(28): p. 5046-5051.
54. SINGH, V.R., *LNESLESS IN-LINE DIGITAL HOLOGRAPHY FOR DYNAMIC APPLICATIONS*, in *School of Mechanical and Aerospace Engineering*. 2008, Nanyang Technological University.
55. Singh, V.R., et al., *Characterization of MEMS cantilevers using lensless digital holographic microscope - art. no. 69950F*, in *Optical Micro- and Nanometrology in Microsystems Technology II*, C. Gorecki, A.K. Asundi, and W. Osten, Editors. 2008, Spie-Int Soc Optical Engineering: Bellingham. p. F9950-F9950.
56. Singh, V.R., et al., *Dynamic characterization of MEMS diaphragm using time averaged in-line digital holography*. Optics Communications, 2007. **280**(2): p. 285-290.
57. Asundi, A. and V.R. Singh, *Amplitude and phase analysis in digital dynamic holography*. Optics Letters, 2006. **31**(16): p. 2420-2422.
58. Asundi, A. and V.R. Singh, *Time-averaged in-line digital holographic interferometry for vibration analysis*. Applied Optics, 2006. **45**(11): p. 2391-2395.
59. Asundi, A. and V.R. Singh, *Sectioning of amplitude images in digital holography*. Measurement Science & Technology, 2006. **17**(1): p. 75-78.
60. Singh, V.R., A. Asundi, and M. Jianrnin, *Double exposure time averaged in-line digital holography - art. no. 661630*, in *Optical Measurement Systems for Industrial Inspection V, Pts 1 and 2*, W. Osten, C. Gorecki, and E.L. Novak, Editors. 2007. p. 61630-61630.
61. Darakis, E., et al., *Phase measurement errors due to holographic interferograms compression - art. no. 661612*, in *Optical Measurement Systems for Industrial Inspection V, Pts 1 and 2*, W. Osten, C. Gorecki, and E.L. Novak, Editors. 2007. p. 61612-61612.
62. Singh, V.R. and A. Asundi, *Advances in dynamic metrology using in-line digital holographic interferometry - art. no. 629306*, in *Interferometry XIII: Applications*, E.L. Novak, W. Osten, and C. Gorecki, Editors. 2006. p. 29306-29306.
63. Singh, V.R. and A. Asundi, *In-line digital holography for dynamic metrology of MEMS*. Chinese Optics Letters, 2009. **7**(12): p. 1117-1122.
64. Singh, V.R., G. Hegde, and A.K. Asundi, *Digital in-line holography for dynamic micro-metrology - art. no. 618803*, in *Optical Micro- and Nanometrology in Microsystems Technology*, C. Gorecki, A.K. Asundi, and W. Osten, Editors. 2006. p. 18803-18803.
65. Kreis, T., *Handbook of Holographic Interferometry, Optical and Digital Methods*. 2005: WILEY-VCH Verlag GmbH & Co.KGaA. 47-53.
66. CuChe, E., P. Marquet, and C. Depeursinge, *Simultaneous amplitude-contrast and quantitative phase-contrast microscopy by numerical reconstruction of Fresnel off-axis holograms*. Applied Optics, 1999. **38**(34): p. 6994-7001.
67. CuChe, E., F. Bevilacqua, and C. Depeursinge, *Digital holography for quantitative phase-contrast imaging*. Optics Letters, 1999. **24**(5): p. 291-293.
68. Qu, W.J., et al., *Transmission digital holographic microscopy based on a beam-splitter cube interferometer*. Applied Optics, 2009. **48**(15): p. 2778-2783.
69. Qu, W.J., et al., *Characterization and inspection of microlens array by single cube beam splitter microscopy*. Applied Optics, 2011. **50**(6): p. 886-890.
70. Qu, W.J., et al., *Quasi-physical phase compensation in digital holographic microscopy*. Journal of the Optical Society of America a-Optics Image Science and Vision, 2009. **26**(9): p. 2005-2011.

REFERENCES

71. Qu, W.J., et al., *Quasi-physical phase compensation in digital holographic microscopy* (vol 26, pg 2005, 2009). Journal of the Optical Society of America a-Optics Image Science and Vision, 2010. **27**(7): p. 1660-1660.
72. Qu, W.J., et al., *Microlens characterization by digital holographic microscopy with physical spherical phase compensation*. Applied Optics, 2010. **49**(33): p. 6448-6454.
73. Qu, W.J., et al., *Digital holographic microscopy with physical phase compensation*. Optics Letters, 2009. **34**(8): p. 1276-1278.
74. Ferraro, P., et al., *Compensation of the inherent wave front curvature in digital holographic coherent microscopy for quantitative phase-contrast imaging*. Applied Optics, 2003. **42**(11): p. 1938-1946.
75. Colomb, T., et al., *Automatic procedure for aberration compensation in digital holographic microscopy and applications to specimen shape compensation*. Applied Optics, 2006. **45**(5): p. 851-863.
76. Colomb, T., et al., *Total aberrations compensation in digital holographic microscopy with a reference conjugated hologram*. Optics Express, 2006. **14**(10): p. 4300-4306.
77. Montfort, F., et al., *Submicrometer optical tomography by multiple-wavelength digital holographic microscopy*. Applied Optics, 2006. **45**(32): p. 8209-8217.
78. Colomb, T., et al., *Numerical parametric lens for shifting, magnification, and complete aberration compensation in digital holographic microscopy*. Journal of the Optical Society of America a-Optics Image Science and Vision, 2006. **23**(12): p. 3177-3190.
79. Rappaz, B., et al., *Measurement of the integral refractive index and dynamic cell morphometry of living cells with digital holographic microscopy*. Optics Express, 2005. **13**(23): p. 9361-9373.
80. Liu, C., et al., *All-in-one multifunctional optical microscope with a single holographic measurement*. Optical Engineering, 2008. **47**(8).
81. Charriere, F., et al., *Cell refractive index tomography by digital holographic microscopy*. Optics Letters, 2006. **31**(2): p. 178-180.
82. Mico, V., Z. Zalevsky, and J. Garcia, *Common-path phase-shifting digital holographic microscopy: A way to quantitative phase imaging and superresolution*. Optics Communications, 2008. **281**(17): p. 4273-4281.
83. Carl, D., et al., *Parameter-optimized digital holographic microscope for high-resolution living-cell analysis*. Applied Optics, 2004. **43**(36): p. 6536-6544.
84. Khmaladze, A., M. Kim, and C.M. Lo, *Phase imaging of cells by simultaneous dual-wavelength reflection digital holography*. Optics Express, 2008. **16**(15): p. 10900-10911.
85. Dubois, F., L. Joannes, and J.C. Legros, *Improved three-dimensional imaging with a digital holography microscope with a source of partial spatial coherence*. Applied Optics, 1999. **38**(34): p. 7085-7094.
86. Rappaz B (Rappaz, B., Charriere F (Charriere, Florian), Depeursinge C (Depeursinge, Christian), Magistretti PJ (Magistretti, Pierre J.), Marquet P (Marquet, Pierre), *Simultaneous cell morphometry and refractive index measurement with dual-wavelength digital holographic microscopy and dye-enhanced dispersion of perfusion medium*. OPTICS LETTERS APR 1 2008. **33** (7): p. 744-746
87. Cuhe, E., P. Marquet, and C. Depeursinge, *Spatial filtering for zero-order and twin-image elimination in digital off-axis holography*. Applied Optics, 2000. **39**(23): p. 4070-4075.
88. Kreis, T., *Handbook of Holographic Interferometry*. 2005.
89. 范琦, 数字全息术中的采样和分辨率研究, in 精密仪器及机械. 2007, 西北工业大学.
90. Kreis, T.M., M. Adams, and W.P.O. Juptner. *Methods of digital holography: A comparison*. in *Conference on Optical Inspection and Micromeasurements II*. 1997. Munich, Germany.
91. Montfort, F., et al., *Purely numerical compensation for microscope objective phase curvature in digital holographic microscopy: influence of digital phase mask position*. Journal of the Optical Society of America a-Optics Image Science and Vision, 2006. **23**(11): p. 2944-2953.
92. Garcia-Sucerquia, J., et al., *Digital in-line holographic microscopy*. Applied Optics, 2006. **45**(5): p. 836-850.

REFERENCES

93. Jin, H.Z., et al., *The influence of structural parameters of CCD on the reconstruction image of digital holograms*. Journal of Modern Optics, 2008. **55**(18): p. 2989-3000.
94. Kelly, D.P., et al., *A Practical Guide to Digital Holography and Generalized Sampling - art. no. 707215*, in *Optics and Photonics for Information Processing II*, A.A.S. Awwal, K.M. Iftekharruddin, and B. Javidi, Editors. 2008, Spie-Int Soc Optical Engineering: Bellingham. p. 7215-7215.
95. Stern, A. and B. Javidi, *Analysis of practical sampling and reconstruction from Fresnel fields*. Optical Engineering, 2004. **43**(1): p. 239-250.
96. Picart, P. and J. Leval, *General theoretical formulation of image formation in digital Fresnel holography*. Journal of the Optical Society of America a-Optics Image Science and Vision, 2008. **25**(7): p. 1744-1761.
97. Kelly, D.P., et al., *Resolution limits in practical digital holographic systems*. Optical Engineering, 2009. **48**(9).
98. Charriere, F., et al., *Characterization of microlenses by digital holographic microscopy*. Applied Optics, 2006. **45**(5): p. 829-835.
99. Charriere, F., et al., *Living specimen tomography by digital holographic microscopy: morphometry of testate amoeba*. Optics Express, 2006. **14**(16): p. 7005-7013.
100. Charriere, F., et al., *Shot-noise influence on the reconstructed phase image signal-to-noise ratio in digital holographic microscopy*. Applied Optics, 2006. **45**(29): p. 7667-7673.
101. Charriere, F., et al., *Amplitude point-spread function measurement of high-NA microscope objectives by digital holographic microscopy*. Optics Letters, 2007. **32**(16): p. 2456-2458.
102. Stern, A. and B. Javidi, *Improved-resolution digital holography using the generalized sampling theorem for locally band-limited fields*. Journal of the Optical Society of America a-Optics Image Science and Vision, 2006. **23**(5): p. 1227-1235.
103. Kreis, T., M. Adams, and W. Juptner, *Aperture synthesis in digital holography*, in *Interferometry XI: Techniques and Analysis*, K. Creath and J. Schmit, Editors. 2002. p. 69-76.
104. Massig, J.H., *Digital off-axis holography with a synthetic aperture*. Optics Letters, 2002. **27**(24): p. 2179-2181.
105. Martinez-Leon, L. and B. Javidi, *Improved resolution synthetic aperture holographic imaging - art. no. 67780A*, in *Three-Dimensional Tv, Video, and Display VI*, B. Javidi, F. Okano, and J.Y. Son, Editors. 2007. p. A7780-A7780.
106. Claus, D., *High resolution digital holographic synthetic aperture applied to deformation measurement and extended depth of field method*. Applied Optics, 2010. **49**(16): p. 3187-3198.
107. Gyimesi, F., et al., *Half-magnitude extensions of resolution and field of view in digital holography by scanning and magnification*. Applied Optics, 2009. **48**(31): p. 6026-6034.
108. Di, J.L., et al., *High resolution digital holographic microscopy with a wide field of view based on a synthetic aperture technique and use of linear CCD scanning*. Applied Optics, 2008. **47**(30): p. 5654-5659.
109. Martinez-Leon, L. and B. Javidi, *Synthetic aperture single-exposure on-axis digital holography*. Optics Express, 2008. **16**(1): p. 161-169.
110. Kreis, T. and K. Schluter, *Resolution enhancement by aperture synthesis in digital holography*. Optical Engineering, 2007. **46**(5).
111. Mico, V., et al., *Superresolved phase-shifting Gabor holography by CCD shift*. Journal of Optics a-Pure and Applied Optics, 2009. **11**(12).
112. Alexandrov, S.A., et al., *Synthetic aperture fourier holographic optical microscopy*. Physical Review Letters, 2006. **97**(16).
113. Hillman, T.R., et al., *High-resolution, wide-field object reconstruction with synthetic aperture Fourier holographic optical microscopy*. Optics Express, 2009. **17**(10): p. 7873-7892.
114. Mico, V., et al., *Synthetic aperture superresolution with multiple off-axis holograms*. Journal of the Optical Society of America a-Optics Image Science and Vision, 2006. **23**(12): p. 3162-3170.
115. Yuan, C.J., H.C. Zhai, and H.T. Liu, *Angular multiplexing in pulsed digital holography for aperture synthesis*. Optics Letters, 2008. **33**(20): p. 2356-2358.

REFERENCES

116. Granero, L., et al., *Synthetic aperture superresolved microscopy in digital lensless Fourier holography by time and angular multiplexing of the object information*. APPLIED OPTICS, 2010. **49**(5): p. 845-857.
117. Feng, P., X. Wen, and R. Lu, *Long-working-distance synthetic aperture Fresnel off-axis digital holography*. Optics Express, 2009. **17**(7): p. 5473-5480.
118. Mico, V., et al., *Superresolved imaging in digital holography by superposition of tilted wavefronts*. Applied Optics, 2006. **45**(5): p. 822-828.
119. Mico, V., Z. Zalevsky, and J. Garcia, *Synthetic aperture microscopy using off-axis illumination and polarization coding*. Optics Communications, 2007. **276**(2): p. 209-217.
120. Mico, V., J. Garcia, and Z. Zalevsky, *Axial superresolution by synthetic aperture generation*. Journal of Optics a-Pure and Applied Optics, 2008. **10**(12).
121. Mico, V., et al., *Superresolution digital holographic microscopy for three-dimensional samples*. Optics Express, 2008. **16**(23): p. 19260-19270.
122. Mico, V., Z. Zalevsky, and J. Garcia, *Superresolution optical system by common-path interferometry*. Optics Express, 2006. **14**(12): p. 5168-5177.
123. Granero, L., et al., *Superresolution imaging method using phase-shifting digital lensless Fourier holography*. Optics Express, 2009. **17**(17): p. 15008-15022.
124. Mico, V. and Z. Zalevsky, *Superresolved digital in-line holographic microscopy for high-resolution lensless biological imaging*. Journal of Biomedical Optics, 2010. **15**(4).
125. Granero, L., Z. Zalevsky, and V. Mico, *Single-exposure two-dimensional superresolution in digital holography using a vertical cavity surface-emitting laser source array*. Optics Letters, 2011. **36**(7): p. 1149-1151.
126. Calabuig, A., et al., *Single-exposure super-resolved interferometric microscopy by red-green-blue multiplexing*. Optics Letters, 2011. **36**(6): p. 885-887.
127. Mico, V., et al., *Transverse resolution improvement using rotating-grating time-multiplexing approach*. Journal of the Optical Society of America a-Optics Image Science and Vision, 2008. **25**(5): p. 1115-1129.
128. Chen, X.L. and S.R.J. Brueck, *Imaging interferometric lithography: approaching the resolution limits of optics*. Optics Letters, 1999. **24**(3): p. 124-126.
129. Schwarz, C.J., Y. Kuznetsova, and S.R.J. Brueck, *Imaging interferometric microscopy*. Optics Letters, 2003. **28**(16): p. 1424-1426.
130. Kuznetsova, Y., A. Neumann, and S.R.J. Brueck, *Imaging interferometric microscopy - approaching the linear systems limits of optical resolution*. Optics Express, 2007. **15**(11): p. 6651-6663.
131. Neumann, A., Y. Kuznetsova, and S.R.J. Brueck, *Structured illumination for the extension of imaging interferometric microscopy*. Optics Express, 2008. **16**(10): p. 6785-6793.
132. Kuznetsova, Y., A. Neumann, and S.R.J. Brueck, *Imaging interferometric microscopy*. Journal of the Optical Society of America a-Optics Image Science and Vision, 2008. **25**(3): p. 811-822.
133. Calabuig, A., et al., *Resolution improvement by single-exposure superresolved interferometric microscopy with a monochrome sensor*. Journal of the Optical Society of America a-Optics Image Science and Vision, 2011. **28**(11): p. 2346-2358.
134. Paturzo, M., et al., *Super-resolution in digital holography by a two-dimensional dynamic phase grating*. Optics Express, 2008. **16**(21): p. 17107-17118.
135. Liu, C., et al., *Super-resolution digital holographic imaging method*. Applied Physics Letters, 2002. **81**(17): p. 3143-3145.
136. Binet, R., J. Colineau, and J.C. Leheureau, *Short-range synthetic aperture imaging at 633 nm by digital holography*. Applied Optics, 2002. **41**(23): p. 4775-4782.
137. Zhang, Y., et al., *Synthetic aperture digital holography by movement of object*, in *Holography, Diffractive Optics, and Applications II, Pts 1 and 2*, Y.L. Sheng, et al., Editors. 2005. p. 581-588.
138. Nakatsuji, T. and K. Matsushima, *Free-viewpoint images captured using phase-shifting synthetic aperture digital holography*. Applied Optics, 2008. **47**(19): p. D136-D143.
139. Colomb, T., et al., *Digital holographic reflectometry*. Optics Express, 2010. **18**(4): p. 3719-3731.

REFERENCES

- 140. Mann, C.J., et al., *High-resolution quantitative phase-contrast microscopy by digital holography*. Optics Express, 2005. **13**(22): p. 8693-8698.
- 141. Wagner, C., et al., *Digital recording and numerical reconstruction of lensless Fourier holograms in optical metrology*. Applied Optics, 1999. **38**(22): p. 4812-4820.
- 142. Picart, P., et al., *Some opportunities for vibration analysis with time averaging in digital Fresnel holography*. Applied Optics, 2005. **44**(3): p. 337-343.
- 143. Le Clerc, F., M. Gross, and L. Collot, *Synthetic-aperture experiment in the visible with on-axis digital heterodyne holography*. Optics Letters, 2001. **26**(20): p. 1550-1552.
- 144. Kreis, T.M., *Frequency analysis of digital holography*. Optical Engineering, 2002. **41**(4): p. 771-778.
- 145. Kreis, T.M., *Frequency analysis of digital holography with reconstruction by convolution*. Optical Engineering, 2002. **41**(8): p. 1829-1839.
- 146. Yan Hao, A.A., *Resolution Analysisi of A Digital Holography System*. Applied Optics, 2011. **50**(2): p. 11.