Paraxial diffraction calculations and their applications in digital optical systems

Dissertation zur Erlangung des
akademischen Grades Doktor-Ingenieur (Dr.-Ing.)

vorgelegt der
Fakultät für Maschinenbau der
Technischen Universität Ilmenau

von M.Sc. Yang Wu

1. Gutachter: Jun.-Prof. Dr. Eng. D. Kelly
2. Gutachter: Prof. Dr. rer. nat. habil. S. Sinzinger
3. Gutachter: Prof. Dr. J. Sheridan

Tag der Einreichung: 10.05.2016
Tag der wissenschaftlichen Aussprache: 12.10.2016

urn:nbn:de:gbv:ilm1-2016000588

Abstract

The paraxial diffraction integral is used to calculate the propagation of the light in wave optics. In this thesis, we first focus on how to calculate the diffraction integral especially the Fresnel transform numerically and semi-analytically. To perform the Fresnel transform, we need know both the amplitude and the phase of the optical field. With the help of digital holography, we are able to measure the phase of the optical field, and thus build a lensless holographic imaging system. The resolution limits of this lensless imaging system are analyzed and experimentally demonstrated. After that a metrology application with this holographic system is introduced and discussed. Phase retrieval is another technique for estimating the phase of an optical field. We examine a novel hybrid phase retrieval-holographic optical system for recovering the phase accurately and robustly. At the end of the thesis, the propagation of partially coherent beams is investigated. The propagation of partially coherent beam is characterized by the 4D mutual coherence integral. Instead of solving this 4D integration, we propose an efficient method for modeling the propagation of partially coherent light. We use this method to examine the effectiveness of DOEs when illuminated by partially coherent light. We find a relationship between the width of the coherence function of the partially coherent source and the feature size of the DOE.

Kurzzusammenfassung

Das paraxiale Beugungsintegral wird in der Wellenoptik verwendet, um die Ausbreitung des Lichts zu berechnen. In dieser Arbeit konzentrieren wir uns zunächst darauf, wie das Beugungsintegral, insbesondere die Fresnel-Transformation numerisch und semi-analytisch gerechnet wird. Zur Durchführung der Fresnel-Transformation müssen wir sowohl die Amplitude als auch die Phase des optischen Feldes kennen. Mit Hilfe der digitalen Holografie sind wir in der Lage die Phase des optischen Feldes messen zu können, und somit ein linsenloses holographisches Abbildungssystem aufzubauen. Die Auflösungsgrenzen dieses linsenlosen Abbildungssystems werden analysiert und experimentell demonstriert. Danach wird eine Messanwendung mit diesem holographischen System eingeführt und diskutiert. Phase Retrieval ist eine weitere Technik zur zu schätzen der Phase. Wir untersuchen ein neuartiges Hybrid-Phase Retrieval-holographisches optisches System um die Phrase genau und robust zu bekommen. Am Ende der Arbeit wird die Ausbreitung von teilweise kohärenten Strahl untersucht. Die Ausbreitung von teilweise kohärenten Strahl lässt sich durch das “4D gegenseitige Kohärenz Integral” beschreiben. Anstelle der Lösung des 4D Integrats, haben wir eine effiziente Methode vorgeschlagen, um die Ausbreitung des teilweisen kohärenten Strahls zu berechnen. Wir verwenden diese Methode, um die Wirksamkeit von DOE zu untersuchen, wenn sie durch teilweise kohärentem Licht beleuchtet. Wir finden eine Beziehung zwischen der Breite der Kohärenzfunktion der teilweise kohärenten Lichtquelle und der Pixelgröße des DOEs.

Contents

Abbreviations and mathematical notation

Symbols	Names
CCD	Charge-coupled device
CMOS	Complementary metal-oxide-semiconductor
DFT	Discrete Fourier transform
ENZ	Extended Nijboer Zernike theory
FFT	Fast Fourier transform
HDRH	Hign dynamical range holography
OFSF	Off-axis Fourier spatial filtering
PSF	Point spread functions
PSI	Phase Shifting Interferometry
RSD	Relative standard deviation operation
SD	Standard deviation operation
SENZ	Simplified Extended Nijboer Zernike theory
SNR	Signal noise rate
a	Aperture radius
$A_{x_1,x_2,x_3}^{x_4,x_5,x_6}$	Clebsch-Gordan coeffcient
$A_f(\rho)$	Algebraic factor in ENZ theory
A_o	Complex object wave at the camera plane
A_r	Complex reference wave at the camera plane
A_{res}	The residual term of the hologram
c_g	Thickness of the object
C, S	Real and imaginary parts of the optical field
d	Piezo motor step size
d_i	Diameter of the pinhole
d_f	Defocus term
D	Half width of the CCD/CMOS sensor
E	Error
E_f	ENZ fitting error
E_N	Numerical error
f	Focal length

$F(\rho)$	Focal factor in ENZ theory
$\mathscr{F}, \mathscr{F}^{-1}$	Forward and inverse Fourier transform operations
FFT, FFT^{-1}	Forward and inverse fast Fourier transforms
FST	Fresnel transform operation
H_n	Captured hologram
H_{AC}, H_{DC}, H_K	Interfere term, DC term and the interfere factor of the hologram
i	Imaginary unit
I, I_n	Intensity of the optical field
I_o^c	Intensity as calculated from PSI capture
I_o^m	Measured intensity when reference wave is blocked
I_s	Intensity profile of the incoherent light source
J_n	nth-order of the Bessel functions of the first kind
k	Wave number
l_t	Spatial coherence width
L	Number of output planes that need to be calculted
L_o	Optical path length
n_g	Refraction index of the object
NA	Numerical aperture
N_n	Electronic sensor noise
$P(\rho, \phi)$	Aberration term
P_n	Legendre polynomial
$P_p(x, y)$	Pupil function
$\Delta P, \Delta Q$	Spatial distance between the tho pinholes in x and y directions
r, ϕ	polar coordinates of input plane
$\Delta r, \Delta\phi$	Sampling intervals in polar coordinate system
r'	Radius of the incoherent light source
R, θ	polar coordinates of output plane
R_n^m	Zernike radial polynomials
t	Time
u_o, v_o	Normalised optical coordinates
u_n, v_n	First and second Lommel functions
u	Complex amplitude of input plane
$\bar{u}$	The Fourier transform of u
u_s	Complex amplitude of the incoherent light source
U	Complex amplitude of output plane
U_N	Numerical solution of the complex amplitude of output plane

U_1, U_2, U_3, U_4	Four different semi-analytical solutions of the focal region
w_n	nth-order of Jinc function
x, y	Cartesian coordinates of input plane
x_n, y_n	Discrete Cartesian coordinates of input plane
$\Delta x, \Delta y$	Sampling intervals in x- and y-directions
X, Y	Cartesian coordinates of output plane
X_n, Y_n	Discrete Cartesian coordinates of output plane
$\Delta X, \Delta Y$	Sampling intervals in X- and Y-directions
z	Distance between input and output plane
Z_n^m	Zernike polynomials
β_n^m	Complex-valued Zernike coefficients
γ	Width of the active area of a pixel
λ	Wavelength
Γ	Mutual coherence function
ψ	Initial phase difference between the object and reference wave
$\Delta\psi$	Phase modulation from the object
ψ_n^s	Phase shift by the piezo motor
ψ_v	Phase change due axial vibration
σ	Standard deviation of the Gaussian function
ρ	Normalised radius of input plane
ρ_c	Correlation cofficient
τ	Exposure time
μ	Complex degree of the coherence
δ	Dirac delta function
δ_T	Dirac comb function

1 Overview

Light, is an electromagnetic wave, which can carry different kinds of information. Its physical parameters, like intensity, phase, frequency, coherence, polarization can be applied to transmit information in different systems. Before photoelectric sensors were invented, film was used to record the intensity of the optical field. On the one hand, it can record the continuous light field without electronic noise, however, it is expensive and inconvenient. With the rapid development of science and technology, the photoelectric sensors like charge-coupled device (CCD) or complementary metal-oxide-semiconductor (CMOS) are now commonly used instead of film. Unlike film, the captured digital intensity field can be easily processed later by computer, and the cost of capturing the light intensity is significant reduced. However the captured field is no longer continuous, but in discrete form, and the spatial sampling rate of the discrete digital signal depends on the pixel size of the sensors and the spacing between them. The technology to process, transport, and storage such optical digital information is referred to as digital optics [1].

In this thesis, several problems in digital optics are investigated. Instead of the ray tracing technique, the diffraction integral is usually used to calculate the propagation of the light in digital optical systems. The paraxial Fresnel diffraction integral is one of the mostly widely used diffraction models. We review its numerical solution in Chap. 2 in this thesis. The uniform sampling of the diffraction integral is a common way to numerically calculate the integration. The input field can be sampled in either Cartesian or cylindrical coordinate systems. The properties of these two sampling techniques are fundamentally different. These two approaches are analyzed and discussed.

Despite the convenient nature of the numerical solution, it also has some disadvantages like aliasing and replicas. If we instead use analytical forms to solve the integration, these disadvantages vanish. However in most cases there are no closed form for the diffraction integration. In Chap. 3, we discuss some special cases with analytic solution of the Fresnel diffraction integral. However analytic solution only exsist for particular input fields like perfect converging wave front, and are not suitable in the general case. Then we introduce the Extended Nijboer Zernike (ENZ) theory, which is a semi-analytical solution of the calculating the optical field near the focus points for general cases by decomposing the input field

into finite terms of Zernike polynomials. As a semi-analytical solution, the ENZ theory gives a very accurate solution to the diffraction integral, however it is very complex to implement. Therefore we we propose the simplified ENZ theory, which reduces the complexity of the ENZ, and decreases the computation time for single diffraction plane calculations. After that, we also use the ENZ theory to calculate the Fresnel diffraction integral for optical planes axially displaced from the focal plane.

To describe the optical field, and use the Fresnel transform to calculate the propagation of the light, we need both the intensity and the phase of the optical field. However as mentioned above, only the intensity can be directly measured, the phase information is more difficult to recover. A widely used method, digital holography is introduced and discussed in Chap. 4. It offers a means of measuring optical phase data and they can be used to estimate a three-dimensional surface or for optical thickness measurements. Letting the object wave interfere with a reference wave, the acquired holograms are recorded by digital sensor array, it is called digital hologram. After processing the digital hologram numerically, the optical phase data can be acquired. Digital holography offers a useful way of measuring optical phase data, which can be used for many metrology applications, like quantitative phase-contrast microscopy[2, 3]. The holographic system we use is a lens-less imaging system, its theoretical resolution limits are discussed in [4]. Different kinds of noise sources in the system are analyzed and handled, which helps us to experimentally approach the theoretical resolution limits. At the end of Chap. 4, we use the digital holographic system to estimate the transmittance function of a “thin” lens.

Another common way to measure the optical phase data is to use phase retrieve algorithms, which is introduced in Chap. 5. Usually we measure two or more intensity patterns of the optical field. Then the optical phase data can be calculated/estimated by iterative algorithms: phase retrieve algorithms. However the algorithm often converges to an incorrect solution due the inappropriate initial phase guess. To acquire a good initial phase guess for the iterative process, we combine the digital holography with phase retrieve technique. The new method is introduced and compared to the other traditional holographic techniques.

Note that all the theoretical problems discussed above assume that the light source is totally coherent. The propagation of the coherent wave in free-space and in the optical system can be mathematical described using a diffraction integral approach. However in practice, the partially coherent, even incoherent beams like LED, are difficult to simulate or model using wave optics. Usually we can model the partially coherent beam by using the mutual coherence integration, however it is a 4D integration, and is difficult to solve, In Chap. 6, we propose a concise

method to simulate the partially coherent beam. Instead of solving the 4D mutual coherence integration by definition, we use the spatially coherent random field to simulate the partially coherent beam. By using the proposed method, we calculate the diffraction pattern of a diffractive optical element (DOE) under partially coherent illumination, and experimentally measure the performance of DOE under partially coherent illumination.

Finally, in Chap. 7 we summarize the main results of this thesis and provide an outlook for further research.

2 Paraxial wave propagation

The following content is discussed in this chapter. The paraxial wave model, which is used throughout this thesis, is introduced. Based on this wave optical model, the Fresnel diffraction integral is used to describe the propagation of optical field in free space. It is not possible to find an analytical solution to the Fresnel transform in general. Therefore it is important to examine numerical methods for calculation the diffraction integral. One approach to performing a Fresnel diffraction is solutions is by sampling the input field uniformly. However this uniform sampling of the input field produces infinite copies in addition to the original solution, known as replicas. This replica effect in Cartesian and cylindrical coordinate systems is discussed. Finally, two fast Fourier transform (FFT) based methods are presented and discussed.

2.1 Introduction to the Fresnel transform

The diffraction integral is used to model the propagation of the light, which accounts for the diffraction phenomenon and wave nature of light. In Goodman's treatment, he began the derivation of a diffraction integral with Maxwell's equations [5]. In the scalar optical description of light, it is assumed that the different vectorial component of the electromagnetic field can be treated independently. Goodman thus uses a single scalar wave equation to describe the field instead of vector treatment. With the help of the Helmholtz equation, Green's theorem, and setting two assumptions that the optical field and its first derivative of the vertical direction outside the aperture are zero, Kirchhoff's diffraction formula[6] was derived. However these two assumptions were later proved to be inconsistent, and were modified by Sommerfeld [5]. By using different boundary conditions, he formed the Rayleigh-Sommerfeld I and II diffraction integral. Under the paraxial approximation, these diffraction integrations reduce to the Fresnel diffraction integration.

2.2 Numerical calculation of the Fresnel diffraction integral

In this section, we examine the Fresnel diffraction integral (Fresnel transform). Because only for particular cases, there are analytical solutions to the Fresnel diffraction integral, it is important to develop robust numerical calculation so that general

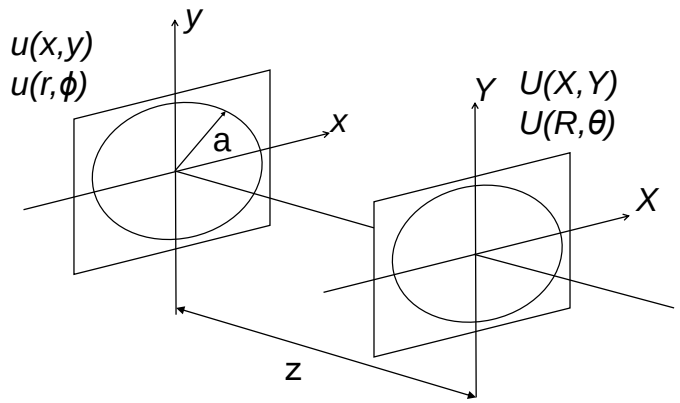


Fig. 2.1: Freespace propagation from input plane $u(x, y)$ to output plane $U(X, Y)$.

diffraction problems can be solved. We discuss a numerical approach where the input field is described using a finite set of samples in both Cartesian and cylindrical coordinate systems. Sampling the input field generates a infinite number of copies of the diffraction field in the output domain. We refer to these copies as replicas and show how they arise in Sec. 2.2.

2.2.1 The Fresnel transform in a Cartesian coordinate system

We begin our analysis by defining the Fresnel transform in same manner as Ref. [5], see Chap. 4, P. 67,

$$\mathrm{FST}\{u(x,y)\} = U(X,Y) = \frac{e^{ikz}}{i\lambda z}\iint_{-\infty}^{+\infty} u\,(x,y)\, e^{\frac{ik}{2z}\left[(X-x)^2+(Y-y)^2\right]} dxdy, \tag{2.1}$$

where FST stands for the Fresnel transform operation, $k = 2\pi/\lambda$ is the wave number, i is the imaginary unit, and λ is the wavelength, $u(x, y)$ is the complex amplitude of the input plane, and $U(X, Y)$ is the complex amplitude of the field in the output plane at a distance of z to the input plane (see Fig. 2.1). As mentioned before, for the most input fields $u(x, y)$, there are no analytical solutions and we must do the integral numerically by sampling $u(x, y)$. In some cases, the input field is already given as discrete values, like lens-less digital holography[7]. If the input field is given continuously, it need to be discretized first, then we have the Fresnel transform in discrete form,

$$U_N(X,Y) = \Delta x \Delta y \frac{e^{ikz}}{i\lambda z} \sum_{n=-\infty}^{\infty}\sum_{m=-\infty}^{\infty} u\,(x_n, y_m)\, e^{\frac{ik}{2z}\left[(X-x_n)^2+(Y-y_m)^2\right]}, \tag{2.2}$$

$$U_N(X,Y) = \Delta x \Delta y \frac{e^{ikz}}{i\lambda z} \iint_S \delta_T(x)\delta_T(y) u(x,y)\, e^{\frac{ik}{2z}\left[(X-x)^2+(Y-y)^2\right]} dxdy, \tag{2.3}$$

$$U_N(X,Y) = \mathrm{FST}\{u(x,y)\delta_T(x)\delta_T(y)\}\Delta x \Delta y, \tag{2.4}$$

where n, m are integers, $x_n = n\Delta x$, $y_m = m\Delta y$, Δx and Δy are the sampling intervals over the x and y coordinate respectively, and $\delta_T(x)$ is the Dirac comb function,

$$\delta_T(x) = \sum_{n=-\infty}^{\infty} \delta(x - n\Delta x), \tag{2.5}$$

where $\delta(x)$ is a Dirac delta function [8],

$$\begin{cases} \delta(x) = +\infty, & x = 0 \\ \delta(x) = 0, & x \neq 0 \\ \int_{+\infty}^{-\infty} \delta(x)dx = 1, \end{cases} \tag{2.6}$$

which we use to define the location of each sample. Then using the Poisson formula for Dirac comb function [9],

$$\sum_{n=-\infty}^{\infty} \delta(x - n\Delta x) = \frac{1}{\Delta x} \sum_{n=-\infty}^{\infty} e^{i2\pi \frac{n}{\Delta x} x}, \tag{2.7}$$

Eq. (2.4) is written as,

$$U_N(X,Y) = \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \mathrm{FST}\left\{u(x,y) e^{i2\pi\left(\frac{nx}{\Delta x} + \frac{my}{\Delta y}\right)}\right\}. \tag{2.8}$$

Note that Eq. (2.8) can also be treated as an infinite sum of Fresnel transforms, by using the shift property of Fresnel transform [10], i.e,

$$\mathrm{FST}\left\{u(x) e^{i2\pi \frac{nx}{\Delta x}}\right\} = U\left(X - \frac{nz\lambda}{\Delta x}\right) \exp\left(\frac{i\pi\lambda z n^2}{\Delta x^2}\right) \exp\left(\frac{i2\pi nx}{\Delta x}\right), \tag{2.9}$$

we have,

$$U_N(X,Y) = \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} U\left(X - \frac{nz\lambda}{\Delta x}, Y - \frac{mz\lambda}{\Delta y}\right) \exp\left(\frac{i\pi\lambda z n^2}{\Delta x^2}\right) \exp\left(\frac{i\pi\lambda z m^2}{\Delta y^2}\right) \times \exp\left(\frac{i2\pi nx}{\Delta x}\right) \exp\left(\frac{i2\pi my}{\Delta y}\right). \tag{2.10}$$

From Eq.(2.10) we find, the numerical approach $U_N(X,Y)$ gives an infinite set orders of replicas centered around the zero order($m=0, n=0$), with the separation $\frac{z\lambda}{\Delta x}$ in the x-direction and $\frac{z\lambda}{\Delta y}$ in the y-direction. Every order of the replicas has the same amplitude, however the phase is different being modulated by $\exp(\frac{i2\pi nx}{\Delta x})\exp(\frac{i2\pi my}{\Delta y})$ in Eq.(2.10). The sampling interval Δx and Δy play an important role in determining the separation between the replicas. The smaller these intervals are, the further away are the replicas from each other in the output plane. On the other hand, if the sampling interval is too large, there will be an overlap between neighboring replicas and the desired zero order solution. This is referred to as aliasing and reduces the accuracy of the numerical solution.

2.2.2 The Fresnel transform in a cylindrical coordinate system

In this section, we will investigate the effects of sampling in a cylindrical coordinate system. First, with the following relations,

$$x = r\cos\phi, \tag{2.11}$$

$$y = r\sin\phi, \tag{2.12}$$

$$X = R\cos\theta, \tag{2.13}$$

$$Y = R\sin\theta, \tag{2.14}$$

we rewrite the Fresnel transform in a cylindrical coordinate system as,

$$U(R,\theta) = \frac{e^{ikz}}{i\lambda z}\iint_S u(r,\phi)\, e^{\frac{ik}{2z}\left(r^2+R^2\right)} e^{-\frac{ikrR}{z}\cos(\phi-\theta)} r dr d\phi, \tag{2.15}$$

In order to solve the integral numerically, we sample the angle and radius with $\Delta\phi$, Δr, then we have,

$$U_N(R,\theta) = \Delta\phi\Delta r\frac{e^{ikz}}{i\lambda z}\times\sum_{n=-\infty}^{\infty}\sum_{m=1}^{M} u(r_n,\phi_m)\, e^{\frac{ik}{2z}\left(r_n^2+R^2\right)} e^{-\frac{ikr_nR}{z}\cos(\phi_m-\theta)} r, \tag{2.16}$$

where $M = 2\pi/\Delta\phi$, $\phi_m = m\Delta\phi$ and $r_n = n\Delta r$.

If we assume the input field $u(r_n,\phi_m)$ is totally symmetric, $u(r_n,\phi_m) = u(r_n)$, we can simplify Eq. (2.16) as,

$$U_N(R) = \Delta r\frac{e^{ikz}}{i\lambda z}\sum_{n=-\infty}^{\infty} u(r_n)\, e^{\frac{ik}{2z}\left(r_n^2+R^2\right)} J_0\left(\frac{2\pi Rr_n}{\lambda z}\right) r_n, \tag{2.17}$$

$$U_N(R) = \Delta r \frac{e^{ikz}}{i\lambda z} \int_0^\infty \delta_T(r) u(r)\, e^{\frac{ik}{2z}\left(r^2+R^2\right)} J_0\left(\frac{2\pi Rr}{\lambda z}\right) r dr, \tag{2.18}$$

where J_0 is the Bessel function of the first kind. Again, with help of the Poisson formula for Dirac comb function, we have,

$$U_N(R) = \frac{e^{ikz}}{i\lambda z} \sum_{n=-\infty}^{\infty} \int_0^\infty e^{i2\pi \frac{nr}{\Delta r}} u(r)\, e^{\frac{ik}{2z}\left(r^2+R^2\right)} J_0\left(\frac{2\pi Rr}{\lambda z}\right) r dr, \tag{2.19}$$

$$U_N(R) = \frac{e^{ikz}}{i\lambda z} \sum_{n=-\infty}^{\infty} H_0\{e^{i2\pi \frac{nr}{\Delta r}} u(r)\, e^{\frac{ik}{2z}\left(r^2+R^2\right)}\}\{2\pi R/\lambda z\}, \tag{2.20}$$

where H_0 is the first order of the Hankel transform[11]. Using the convolution property of Hankel transform, see Chap. 17 of Ref. [11]),

$$H_0\{e^{i2\pi \frac{nr}{\Delta r}} u(r)\, e^{\frac{ik}{2z}\left(r^2+R^2\right)}\}\{2\pi R/\lambda z\} = \frac{1}{2\pi} H_0\{e^{i2\pi \frac{nr}{\Delta r}}\} * H_0\{u(r)\, e^{\frac{ik}{2z}\left(r^2+R^2\right)}\}, \tag{2.21}$$

where "$*$" is the convolution operation,

$$(f * g)(X) = \int_\infty^\infty f(x) g(X-x) dx. \tag{2.22}$$

Note that the analytical expression of $U(R)$ can be also written as a Hankel transform,

$$U(R) = \frac{e^{ikz}}{i\lambda z} \int_0^\infty u(r)\, e^{\frac{ik}{2z}\left(r^2+R^2\right)} J_0\left(\frac{2\pi Rr}{\lambda z}\right) r dr = H_0\{u(r)\, e^{\frac{ik}{2z}\left(r^2+R^2\right)}\}. \tag{2.23}$$

Furthermore we know that the Hankel transform of an exponential function $e^{-\beta r}$, is given by,

$$H_0\{e^{-\beta r}\}\{v\} = \frac{\beta}{(v^2+\beta^2)^{\frac{3}{2}}}. \tag{2.24}$$

Substituting Eq. (2.21), Eq. (2.23) and Eq. (2.24) into Eq. (2.20), we arrive of the following expression,

$$U_N(R) = \sum_{n=-\infty}^{\infty} U(R) * \frac{-i2\pi \frac{n}{\Delta r}}{\left[\left(\frac{2\pi R}{\lambda z}\right)^2 - (2\pi \frac{n}{\Delta r})^2\right]^{\frac{3}{2}}}. \tag{2.25}$$

We see from the equation above, the fraction term on the right side acts like a Delta function at different position $R = |n\lambda z/\Delta r|$, which means the sampling over the radius also produces an infinite number of replicas in the output domain. However these replicas are fundamentally different from the replicas in Cartesian coordinate system. The higher order replicas here are not identical to the zero order replica, they are located at distances of $R = |n\lambda z/\Delta r|$. Therefore the positive order $+n$ and negative order $-n$ replicas overlap with each other.

We now focus on the problem of numerically calculating the integration over the angle ϕ with a finite number of samples as in Eq. (2.15). To make the analysis easier, we examine the integral over the angle separately,

$$U(R,\theta) = A(\theta)\frac{e^{ikz}}{i\lambda z}\int_0^{\infty} u\,(r)\, e^{\frac{ik}{2z}\left(r^2+R^2\right)} r dr, \tag{2.26}$$

$$A(\theta) = \int_0^{2\pi} e^{-\frac{ikrR}{z}\cos(\theta-\phi)} d\phi. \tag{2.27}$$

After we sample the angle variable ϕ, we have the following numerical expression,

$$A_N(\theta) = \Delta\phi \sum_{m=1}^{M} e^{-\frac{ikrR}{z}\cos(\theta-\phi_m)} \tag{2.28}$$

$$= \Delta\phi \int_0^{2\pi} \delta_T(\phi_m) e^{-\frac{ikrR}{z}\cos(\theta-\phi)} d\phi \tag{2.29}$$

$$= \sum_{m=-\infty}^{\infty} \int_0^{2\pi} e^{imM\phi} e^{-\frac{ikrR}{z}\cos(\theta-\phi)} d\phi \tag{2.30}$$

which $A_N(\theta)$ is the sampled form of $A(\theta)$. Substituting Eq. (2.31), which is given in Chap. 2.2, Eq. (5) in Ref. [12],

$$J_n(\alpha) = -\frac{1}{2\pi}\int_0^{2\pi} e^{i[n\phi-\alpha\cos(\phi)]} d\phi, \tag{2.31}$$

into Eq. (2.28), we arrive at the following result,

$$A_N(\theta) = \sum_{m=-\infty}^{\infty} e^{imM\phi} 2\pi J_{mM}\left(\frac{krR}{z}\right) = A(\theta) + E_N, \tag{2.32}$$

which E_N is the numerical error with,

$$E_N = \sum_{m=1}^{\infty} 2\pi J_{mM}(v) \left[e^{imM\phi} + (-1)^M e^{-imM\phi} \right]. \tag{2.33}$$

The accuracy of the angular sampling depends on this error E_N. It is a summation over the different orders of Bessel function J_M, J_{2M}, J_{3M}.... If the term $v = \frac{krR}{z}$ is small $(0 < v < 1)$, the contribution from the higher orders of Bessel functions are approximately zero,

$$E_N \approx 2\pi J_M(v) \left[e^{iM\phi} + (-1)^M e^{-iM\phi} \right]. \tag{2.34}$$

In this case, when $M > 5$, $J_M(v) \approx 0$ and therefore $E_N \approx 0$, the sampling will give a reasonable accurate result. However if we want to calculate the points far away from the optical axis, which means $v >> 1$, a higher number of sampling points M are needed, to ensure numerically accurate results.

2.3 Some numerical examples

In order to verify our mathematical analysis above, we calculate some examples of Fresnel transform numerically, by sampling the input field in the two different coordinate systems, and examine the properties of the resulting replicas.

2.3.1 A symmetrical converging spherical wave

In this Cartesian case, we want to calculate the optical field in the output plane when a perfect converging spherical wave is incident upon a circular aperture. We choose the aperture radius $a = 0.25$ mm, wavelength $\lambda = 500$ nm, the focal point of the converging wave is $f = 50$ mm and the sampling interval is $\Delta x = \Delta y = 0.025$ mm. The input field in Cartesian coordinate system is then given in Ref. [5] in Chap. 5,

$$\begin{cases} u(x,y) = e^{-\frac{i\pi}{\lambda f}(x^2+y^2)}, & x^2 + y^2 \leq a^2 \\ u(x,y) = 0, & x^2 + y^2 > a^2 \end{cases} \tag{2.35}$$

Substituting the input field $u(x, y)$ into Eq. (2.2), the amplitude distribution at the focus plane $z = 50$ mm is calculated and the result plotted in Fig. 2.2. From Eq. (2.10) we know that an infinite numbers of replicas are generated along with the original solution. These replicas are separated from each other by a distance of $\lambda z/\Delta x = 1$ mm, which is in accordance with the simulation result in Fig. 2.2. We not that if the Δx and Δy are too big, these replicas would overlap with the original result, producing aliasing and thus reducing the accuracy of the result.

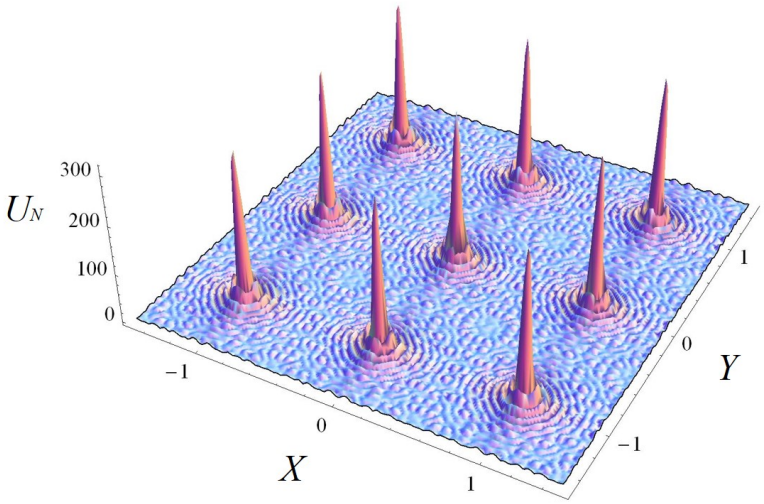


Fig. 2.2: Amplitude distribution of a converging spherical wave at the focus plane, see details in text. The input field is sampled in Cartesian coordinate system, with $\Delta x = \Delta y = 0.025$ mm. Therefore the replicas are seen at the distance of $\lambda z/\Delta x = 1$ mm.

We now repeat the calculation for the same input field however this time using a cylindrical coordinate system. In this case the cylindrical input field is written as,

$$\begin{cases} u(r,\theta) = e^{-\frac{i\pi}{\lambda f}r^2}, & r \le a \\ u(r,\theta) = 0, & r > a \end{cases} \tag{2.36}$$

In order to investigate this calculation in more detail we sample over the radius r and the angle θ separately. We first look at the replica along the output radial coordinate. We sampled the radius with $\Delta r = 0.025$ mm, then substituting Eq. (2.36) into Eq. (2.17). The amplitude is plotted in Fig. 2.3. As predicted by Eq. (2.25), we can see replicas with the distance of $R = \lambda z/\Delta r = 1$ mm. However unlike the replicas in Fig. 2.2, where the amplitudes of different orders of replicas are identical to the original one, the replicas here are totally different to the original result. In fact the positive and negative orders overlap at the same location.

If we sample the angle in cylindrical coordinate system using Eq. (2.28), the error depends on the sampling number M and calculated radius $v = krR/z$ in output plane. In order to validate Eq. (2.34), we use Eq. (2.28) to calculate Eq. (2.27) with $v_o = krR/z = 1$ and $M = 10$. The result is plotted in Fig. 2.4. According to Eq.

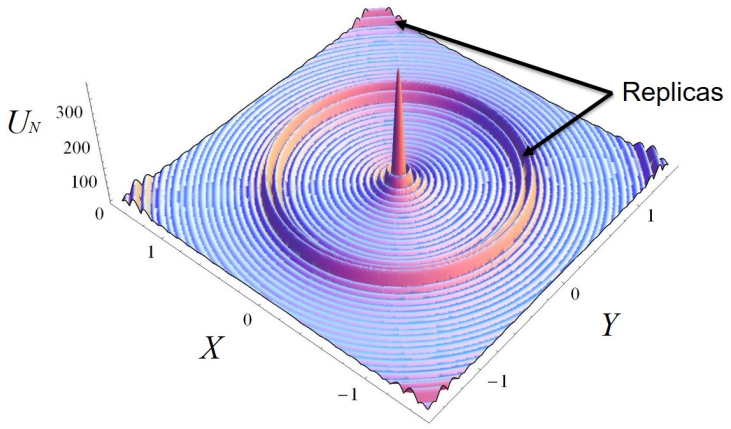


Fig. 2.3: Amplitude distribution of a converging spherical wave at the focus plane. The input field is sampled in cylindric coordinate system, with $\Delta r = 0.025$ mm. Therefore the replicas are seen at the distance of $\lambda z/\Delta r = 1$ mm.

(2.34), the numerical error should be,

$$E_N \approx 4\pi J_{10}(1)\cos(10\phi) = 3.3057 \times 10^{-9}\cos(10\phi), \tag{2.37}$$

which matches the result in Fig. 2.4, there are 10 periods in $0 - 2\pi$. The bigger v_o is, the more sampling points M are needed to get an accurate result.

2.3.2 An asymmetrical input field

In the previous section, we analyzed and calculated the diffracted field when a completely symmetrical input field was used in both Cartesian and cylindrical coordinate systems. In this section, we will calculate the asymmetrical input field by sampling in cylindrical coordinate systems.

We take the shifted converging field as our input field , instead Eq. (2.36). We have now the asymmetrical input field in cylindrical coordinate systems with,

$$u(r,\phi) = e^{-\frac{i\pi}{\lambda f}(r^2)} e^{i2\pi r\cos(\phi)/0.1}. \tag{2.38}$$

The focal point is therefore shifted and located at $x = 0.25$ mm. As a reference we calculate the results by sampling the input field in Cartesian coordinate system, with very fine sampling $\Delta x = \Delta y = 0.001$ mm, to make sure the replicas do not affect the accuracy of the results. The amplitude distribution is plotted in Fig. 2.5. In Fig. 2.6-2.8, the amplitude distribution, by sampling in a cylindrical coordinate system, are plotted, with same sampling interval of radius $\Delta r = 0.05$ mm, but different

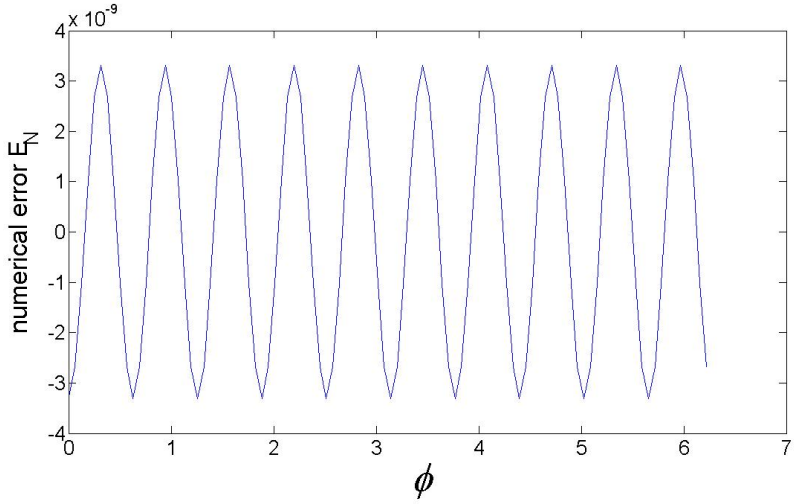


Fig. 2.4: Numerical error E_N by sampling the angle with $M = 10$ and $v_o = 1$.

sampling number M over ϕ. Therefore we can see replicas in Fig. 2.6-2.8, with the same replicas distances, which is given by $\lambda z/\Delta r = 0.5$ mm. In Fig. 2.6, $M = 300$ is taken to uniformly sample the angle ϕ, replicas are averagely distributed around the focus point. When we reduce the sampling number M to 80, as shown in Fig. 2.7, the error are seen at the third order of replica. If less sampling number $M = 40$ is taken, which is shown in Fig. 2.8, the error increases and even the first order of replica is effected by the undersampling of ϕ. The further away from the zero order the observation location is, the more samples M are required to keep the error within given bounds.

2.4 FFT-based calculation

Regardless of which coordinate system is chosen, sampling the input field and doing the integral directly will take a large calculation time. For example for an input field is sampled with 200×200 points, to calculate the complex amplitude at a single output location takes about 0.04 s with the following computer: Intel(R) Core(TM) i7-2600K. Therefore the calculation of an output field with 200×200 points will take about 25 minutes. In this section, we discuss the numerical solution of the Fresnel transform using the highly efficient FFT-based calculation, which can dramatically reduce the computation time.

The Fourier transform is a linear integral transformation [13]. By transforming the signal between the time domain and frequency domain, it has many applications in physics and engineering. In digital optics, we use it to transform the optical signal between the spatial domain and spatial-frequency domain, its applications are

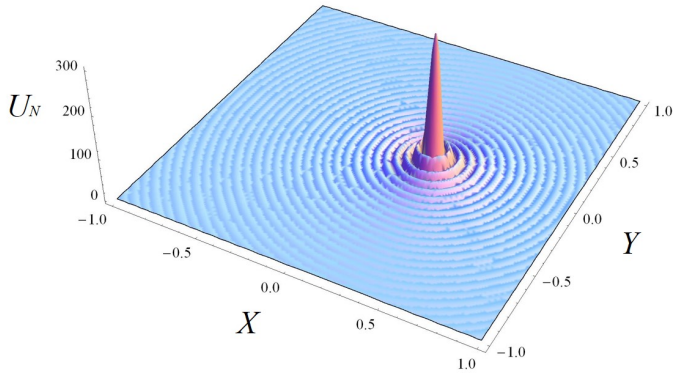


Fig. 2.5: Amplitude contour $|U_N(X, Y)|$ by sampling in the Cartesian coordinate system with $\Delta x = \Delta y = 0.001$ mm. Here the replicas are far enough apart from each other, to produce accurate numerical results.

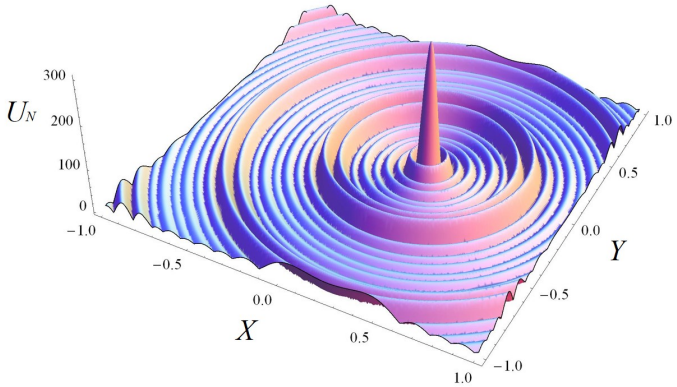


Fig. 2.6: Amplitude contour $|U_N(X, Y)|$ by sampling in the cylindrical coordinate system with $\Delta r = 0.05$ mm, $M = 300$. Here the circle replicas are seen at a distance of $\lambda z/\Delta r = 0.5$ mm. Because the angle is very fine sampled, the numerical error due the angle sampling is very small.

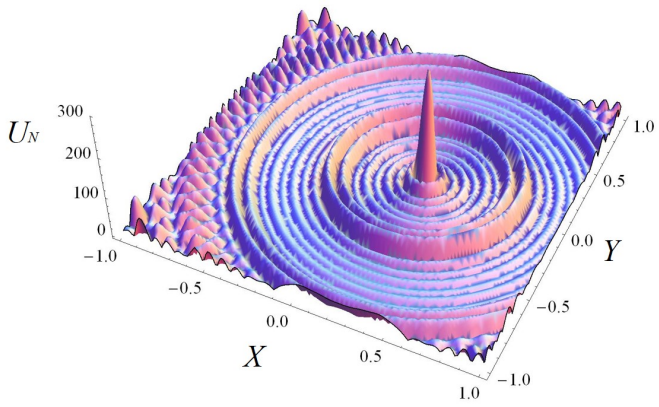


Fig. 2.7: Amplitude contour $|U_N(X, Y)|$ by sampling in the cylindrical coordinate system with $\Delta r = 0.05$ mm, $M = 80$. The angle is sampled with $M = 80$, therefore we can see the error due the angle sampling appears at the position, where v_o is large.

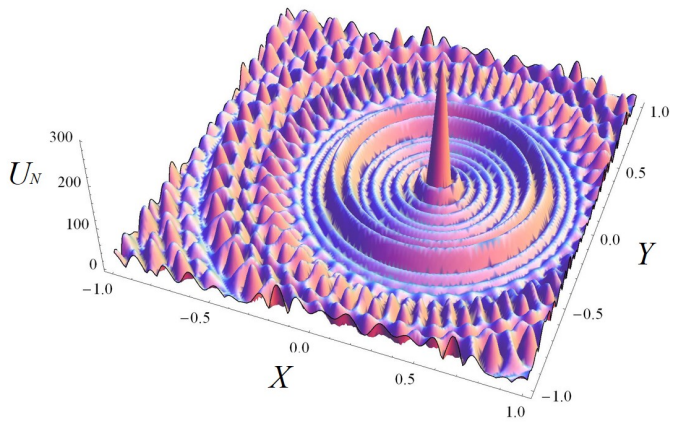


Fig. 2.8: Amplitude contour $|U_N(X, Y)|$ by sampling in the cylindrical coordinate system with $\Delta r = 0.05$ mm, $M = 40$. In this case, the angle sampling is not enough, the error can be seen even near centre of the diffraction plane.

found in diffraction, imaging, optical data processing, and holography. The Fourier transform of a continuous function is,

$$\bar{u}(X) = \mathscr{F}\{u(x)\}\{X\} = \int u(x)e^{-i2\pi xX}dx, \tag{2.39}$$

and the inverse Fourier transform is given by,

$$u(x) = \mathscr{F}^{-1}\{\bar{u}(X)\}\{x\} = \int \bar{u}(X)e^{i2\pi xX}dX, \tag{2.40}$$

where $\mathscr{F}$ and $\mathscr{F}^{-1}$ are the forward and inverse Fourier transform operations. For the 2D case, the Fourier transform is then,

$$\bar{u}(X,Y) = \mathscr{F}\{u(x,y)\}\{X,Y\} = \iint u(x,y)e^{-i2\pi xX}e^{-i2\pi yY}dxdy. \tag{2.41}$$

Again it is not possible to solve the Fourier integral for general input function. We must again turn to numerical approaches. In a discrete system, a computer can calculate for a finite number of samples. Therefore we change the Fourier transform to the discrete Fourier transform. We rewrite the Fourier transform Eq. (2.39) in discrete form,

$$\bar{u}(X) = \Delta x \sum_{n=1}^{N} u(x_n)e^{-i2\pi x_n X}, \tag{2.42}$$

which $u(x_n)$ has a sequence with N samples $u(x_1), u(x_2), ..., u(x_N)$, the sampling interval is Δx. According to Eq. (2.42), the calculated output signal $\bar{u}(X)$ is continuous. Calculating the Eq. (2.42) uses a large computational load $O(n^2)$. The use of fast Fourier transform algorithm(FFT) reduces the computational load to $O(n \log n)$ instead. As same as Eq. (2.42), we have a discrete signal $u(x_n)$ with N samples, and the sampling interval Δx. After the FFT operation, we get the same output signal $\bar{u}(X_n)$ however in discrete form,

$$\bar{u}(X_n) = \mathrm{FFT}\left\{u(x_n)\right\}. \tag{2.43}$$

The acquired output signal $\bar{U}(X_n)$ has the same sequence size N as the input signal, however the sampling interval of the output signal is changed with $\Delta X = 1/(N\Delta x)$, and the output signal extent is therefore $1/\Delta x$. Note that the input signal is uniformly sampled, which produces also replica effect similar to the discussion in Sec. 2.2.1. The distance between replicas is $1/\Delta x$. We note that, the sampling interval of the output signal is fixed with ΔX and the extent of the output signal is limited by $N\Delta X$. If we want to calculate the output signal with finer sampling rate and larger extent, zero-padding and interpolation could be used.

Zero-padding

For example a input signal $u = [u(x_1), u(x_2)...u(x_N)]$ that is sampled with N samples and with a sampling interval Δx, we put extra zeros beside the signal, and have the signal u_{zp},

$$u_{zp} = [0, 0, 0, ...u..., 0, 0, 0], \tag{2.44}$$

the vector u_{zp} in Eq. (2.44) has M samples($M > N$), and $M - N$ is the number of the adding zeros, we call this technique zero-padding. In this case, the extent of the output signal does not change, but with more sampling points M, so the sampling interval of the output signal is $1/M\Delta x$ instead of $1/N\Delta x$.

Interpolation

Another technique to manipulate the sampling of the output signal is the interpolation. If we interpolate the input signal u, then we have the signal u_{ip},

$$u_{ip} = [u(x_1), u(x^1), u(x_2), u(x^2), u(x_3)..., u(x_N)], \tag{2.45}$$

which $u(x^n)$ means the interpolated terms between $u(x_n)$ and $u(x_{n+1})$, with

$$u(x^n) = \frac{u(x_n) + u(x_{n+1})}{2}. \tag{2.46}$$

In this instance, the original signal u is replaced into u_{ip} with a smaller sampling interval, which leads to a larger extent of the output signal, however the sampling interval of the output signal is still the same.

By properly combining these two numerical techniques we can manipulate the sampling rate and the extent of the output signal, and calculate the signal for any spatial frequency coordinate we want.

We now return to our discussion and calculate the Fresnel transform numerically. In order to perform the calculation of the Fresnel transform, we rewrite the Fresnel transform into the form of the Fourier transform. By using the FFT algorithms, the speed of calculating the Fresnel transform is significantly improved. Generally there are two methods to calculate the Fresnel transform using FFT technique, the direct method and the spectral method. Next, we will examine these two FFT-based methods.

2.4.1 The direct method

We rewrite the Fresnel transform from Eq. (2.1) as following,

$$\mathrm{FST}\{u(x,y)\} = U(X,Y) = e^{\frac{ik(X^2+Y^2)}{2z}} \frac{e^{ikz}}{i\lambda z} \iint_{-\infty}^{+\infty} u(x,y)\, e^{\frac{ik(x^2+y^2)}{2z}} e^{\frac{-ik(xX+yY)}{z}} dxdy. \tag{2.47}$$

Note that we can rewrite the equation above in the form of 2D Fourier transform, with the scaled factor $v_x = X/(\lambda z), v_y = Y/(\lambda z)$,

$$\mathrm{FST}\{u(x,y)\} = U(X,Y) = e^{\frac{ik(X^2+Y^2)}{2z}} \frac{e^{ikz}}{\lambda z} \mathscr{F}\left\{u(x,y)\, e^{\frac{ik(x^2+y^2)}{2z}}\right\}\{v_x, v_y\}. \tag{2.48}$$

Using FFT to calculate the Fourier transform, we have the discrete output field with direct method,

$$\mathrm{FST}\{u(x_n,y_n)\} = U(X_n,Y_n) = e^{\frac{ik(X_n^2+Y_n^2)}{2z}} \frac{e^{ikz}}{\lambda z} \mathrm{FFT}\left\{u(x_n,y_n)\, e^{\frac{ik(x_n^2+y_n^2)}{2z}}\right\}. \tag{2.49}$$

According to the property of the FFT operation, the discrete output field $U(X_n, Y_n)$ has the same matrix size as the input field $u(x_n, y_n)$. The sampling interval and the extent of the output field is $(\lambda z/N\Delta x, \lambda z/N\Delta y)$ and $(\lambda z/\Delta x, \lambda z/\Delta y)$ in both coordinates.

Note that unless $\lambda z/N = \Delta x^2, \lambda z/N = \Delta y^2$, the extent of output field is different from the extent of input field, the output extent size depends on the factor $\lambda z/N$. Therefore when we calculate the Fresnel transform using direct method at different propagation distances z or for different λ, the sampling interval and the extent of the output field are changed accordingly.

2.4.2 The spectral method

Similar to the direct method, the Fresnel transform is rewritten in the form of Fourier transform. We note that the Fresnel transform can be expressed as a convolution operation,

$$U(X,Y) = \frac{e^{ikz}}{i\lambda z}\left[u(x,y) * e^{\frac{ik(x^2+y^2)}{2z}}\right], \tag{2.50}$$

$$(f*g)(t) = \int_{-\infty}^{\infty} f(\tau)g(t-\tau)d\tau, \tag{2.51}$$

Table 2.1: Comparison between direct method and spectral method

output field property	direct method	spectral method
sampling interval	$\lambda z/N\Delta x$	Δx
output extent	$\lambda z/\Delta x$	$N\Delta x$
sampling interval after zero-padding	$\lambda z/M\Delta x$	Δx
extent after zero-padding	$\lambda z/\Delta x$	$M\Delta x$
sampling interval after interpolation	$\lambda z/N\Delta x$	$N\Delta x/M$
extent after interpolation	$M\lambda z/N\Delta x$	$N\Delta x$

where "*" is the convolution operation. Then make use of Fourier property of convolution,

$$\mathscr{F}\{U(X,Y)\} = \frac{e^{ikz}}{i\lambda z}\mathscr{F}\{u\,(x,y)\}\mathscr{F}\{e^{\frac{ik(x^2+y^2)}{2z}}\}, \tag{2.52}$$

$$U(X,Y) = \frac{e^{ikz}}{i\lambda z}\mathscr{F}^{-1}\left\{\mathscr{F}\{u\,(x,y)\}\{\bar{x},\bar{y}\} \times \mathscr{F}\{e^{\frac{ik(x^2+y^2)}{2z}}\}\{\bar{x},\bar{y}\}\right\}\{X,Y\}\,. \tag{2.53}$$

Similarly we use discrete FFT instead of continuous Fourier transform,

$$U(X_n,Y_n) = \frac{e^{ikz}}{i\lambda z}\mathrm{FFT}^{-1}\left\{\mathrm{FFT}\{u\,(x_n,y_n)\} \times \mathrm{FFT}\{e^{\frac{ik(x_n^2+y_n^2)}{2z}}\}\right\}. \tag{2.54}$$

In the spectral method, two FFTs are used in the convolution. This ensures that the sampling interval of the output field is as same as the sampling interval of the input field, $\Delta X = \Delta x$, $\Delta Y = \Delta y$. Hence the extent of output field is always equal to the extent of input field, no matter what the propagation distance z is.

Note that both methods produce replicas, the distance between the two neighboring replicas equals to the extent of the output field. If the extent of the output field is smaller than the signal width, there would be aliasing in the output field. To separate the replicas we increase the extent of the output field by using the zero-padding or the interpolating techniques, which will be discussed in the next section.

2.4.3 A comparison of the direct and spectral methods

After introducing the two FFT-based techniques, we discuss their properties and some advantages and disadvantages of the two techniques in this section. Assuming the input field $u(x_n, y_n)$ has a matrix of $N \times N$ points, the sampling interval are $\Delta x, \Delta y$ in x and y coordinates, the differences of these two methods are listed in Tab.

2.1.

As mentioned above, the extent and the sampling interval of output field do not change while using spectral method, however by using direct method, the extent and the sampling interval are changed proportional to the propagation distance z. In addition, the differences are apparent when we manipulate the input field by using zero-padding or interpolation. For the direct method, if we zero-pad the input field to a matrix size $M \times M (M > N)$, the extent of output field does not change, but we do get a finer sampled input field with the sampling interval $\lambda z / M\Delta x$ in x-coordinate, and corresponding $\lambda z / M\Delta y$ in y-coordinate. For the spectral method, the effect of the zero-padding is opposite to the direct method, the sampling interval is unchanged, while the extent/distance between two neighboring replicas is expanded to $M\Delta x, M\Delta y$. Another numerical technique for controlling the sampling property in the output domain is by interpolation. After linear interpolation, the input field has a matrix size $M = 2N$ or $3N$ depending on whether we upsample by a factor of two or three. In this case, the sampling interval of output domain does not change while using direct method, but the extent/distance between two neighboring replicas is enlarged to $M\lambda z / N\Delta x$. For the spectral method, the extent remains unchanged, however the sampling interval is reduced to $N\Delta x / M$.

Although FFT-based methods give us the discrete solution of the output field $U(X_n, Y_n)$, by appropriately using the zero-padding and the interpolation techniques any spatial locations in the output field can be calculated with these two methods. However there is still a crucial difference between them. In spectral method, after the first FFT, we change the input field from spatial domain to the spatial-frequency domain. If the input field contains very high spatial frequency component, such high spatial frequency part would cause aliasing in the spatial-frequency domain. Therefore some input fields with high spatial frequency are not suitable for using the spectral method, however for direct method, there is no such problem.

2.5 Conclusion

In this chapter, the Fresnel transform was introduced. To solve the Fresnel transform numerically, we examined how sampling the input field effected the accuracy of the numerical results in two different coordinate systems. We also examined the relationship between the correct analytical solution and higher order replicas that arise due to the sampling process. The use of the FFT algorithm reduces the calculation time of the Fresnel transform significantly. Two different FFT-based methods were demonstrated, and their properties were discussed.

3 A semi-analytical solution of the diffraction integral

In Chap. 2, we introduced the Fresnel transform and its numerical solution. With the direct integral in Sec. 2.2, the calculation is time-consuming. If we rewrite the Fresnel transform in the form of a Fourier transform and use an FFT-based technique to calculate the output field with either spectral method or direct method, the calculation time is significantly reduced. A significant disadvantage however is that the resulting output field is discrete and the sampling interval and the extent of the output plane are determined by the sampling of the input plane. Also the replicas are inevitable if the input field is uniformly sampled.

In most cases there are no closed form for the diffraction integral. Therefore in this chapter, we focus on the semi-analytical solution of the diffraction integral. There are several good reasons for investigating diffraction problems using analytical techniques: (i) more insight into the diffraction process is provided, (ii) a "correct" analytical solution serves as an excellent way of testing the predictions of numerical calculations in specific cases, and (iii) sometimes it may be desirable to use a combination of analytical and numerical techniques to best model a given diffraction problem.

3.1 A focal region calculation for a perfectly converging lens

We assume that the paraxial approximation is valid and also use the Fresnel transform to model the diffraction and propagation of light in free space. As mentioned before, for most cases, there are no analytical solutions for the Fresnel transform. In this section we examine the diffraction of monochromatic light when a perfect converging lens that is limited in extent by a circular aperture. By perfect converging lens we mean there is no abberations in the lens, or in the viewpoint of geometrical optics, all the light rays focus to an single focal point. To calculate this focal region problem analytically, Lommel [14] introduced two functions to solve the integral in 1885. In 1947, Nijboer [15][16] gave another solution of the integral using Zernike polynomials. With this solution we can calculate the 3D field distribution behind a perfect converging lens. This solution can also be extended to include the effects of abberation. In 2002, Cao [17] developed another series expansion, and used it to solve the diffraction integral analytically.

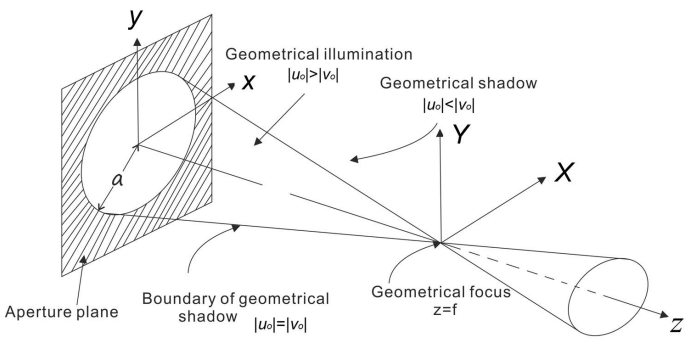


Fig. 3.1: Schematic view of the diffractive geometric.

These semi-analytical solutions are all mathematical equivalent. However none of them are a closed form solution and hence require an infinite sum of series terms. To calculate the result from one of these solutions, in practice, requires truncating the infinite series when a desired accuracy has been achieved. As we shall see each solution, has different convergence properties, which vary depending on the spatial location in the output plane. In this section we find some of these solutions are stable and converge quickly in some regions, however do not perform so well in other locations. We analyse the reasons behind this and provide guidelines on how to choose the appropriate solution for a given spatial location.

We use Fig. 3.1 to illustrate the optical system that we wish to analyze. A plane wave is incident on a perfect converging spherical lens. In the geometrical approximation the focused light would converge to an ideal point source at the focus. In paraxial wave optics however, diffraction introduced by the finite extent of the focusing lens aperture, causes the complex amplitude distribution in the focal plane to spread out over the plane. If one traces the intensity distribution in the focal plane, moving radially out from the focus, one observes a bright central lobe which changes to series of bright and dark rings as one moves out along the plane. We wish to examine the distribution over the entire focal volume, which requires a more complex analytical solution. Using the Fresnel transform in cylindrical coordinate system(Eq. (2.11-2.15)), and substituting the input field of a perfect converging spherical lens(Eq. (2.36)), we have,

$$U(R,\theta) = \frac{e^{ikz}}{i\lambda z} \int_0^a \int_0^{2\pi} e^{-\frac{ik}{2f}(r^2)} e^{\frac{ik}{2z}\left(r^2+R^2\right)} e^{-\frac{ikrR}{z}\cos(\phi-\theta)} r dr d\phi. \tag{3.1}$$

Then using a property of the Bessel function (see Ref. [5], Chap. 2),

$$J_0(\boldsymbol{\alpha}) = \frac{1}{2\pi}\int_0^{2\pi} e^{-i\boldsymbol{\alpha}\cos(\theta-\phi)}d\theta, \tag{3.2}$$

we arrive at the following result,

$$U(R,z) = \frac{2\pi}{i\lambda z}e^{jk\left(z+\frac{R^2}{2z}\right)}\int_0^a e^{\frac{i\pi r^2}{\lambda}\left(\frac{1}{z}-\frac{1}{f}\right)}J_0\left(\frac{2\pi Rr}{\lambda z}\right)rdr. \tag{3.3}$$

After introducing the normalise radius ρ,

$$0 \le \rho = r/a \le 1, \tag{3.4}$$

and using two normalised optical coordinates u_o, v_o to make the integral concise,

$$u_o = \frac{2\pi a^2}{\lambda}\left(\frac{1}{f}-\frac{1}{z}\right), \tag{3.5}$$

$$v_o = \frac{2\pi aR}{\lambda z}, \tag{3.6}$$

we get,

$$U(u_o,v_o) = \int_0^1 e^{-\frac{iu_o\rho^2}{2}}J_0(v_o\rho)\rho d\rho. \tag{3.7}$$

Note that we are interested in the intensity of the output field, therefore the phase term and the amplitude constant outside the integral are ignored. Using Eq. (3.7) we can calculate the complex amplitude of the points behind the lens.

In the next section, however we first concentrate on finding a solution to the integral $U(u_o, v_o)$, using four different semi-analytical methods. As shown in Fig. 3.1, the diffraction field can be divided into two areas: if $|u_o| > |v_o|$, it is in the illumination area; And if $|u_o| < |v_o|$, it is in the geometrical shadow area. Note that, if $|u_o| = |v_o|$, the point is at the boundary between the illumination area and the shadow area, and that the focal point is located at the coordinates $u_o = 0$, $v_o = 0$. It seems reasonable to expect that the various semi-analytical solutions have different properties in each of these regions.

3.1.1 A semi-analytical solution in the focal region

In this section, we derive four different semi-analytical solutions for the diffraction integral, Eq. (3.7).

I. The first Lommel solution

$U(u_o, v_o)$ is separated into real and imaginary parts[14],

$$U(u_o, v_o) = C(u_o, v_o) + iS(u_o, v_o), \tag{3.8}$$

where

$$C(u_o, v_o) = \int_0^1 J_0(v_o\rho) \cos\left(\frac{1}{2}u_o\rho^2\right) \rho d\rho, \tag{3.9}$$

$$S(u_o, v_o) = \int_0^1 J_0(v_o\rho) \sin\left(\frac{1}{2}u_o\rho^2\right) \rho d\rho. \tag{3.10}$$

Using integral by parts,

$$\int A(x)B'(x)dx = A(x)B(x) - \int A'(x)B(x)dx, \tag{3.11}$$

where $A(x) = \cos(\frac{1}{2}u_o\rho^2)$ and $B'(x) = J_0(v_o\rho)$, and using the property of Bessel function (Ref. [12], Page 18),

$$\frac{d}{dx}\left[x^{n+1}J_{n+1}(x)\right] = x^{n+1}J_n(x). \tag{3.12}$$

we get that,

$$C(u_o, v_o) = \frac{1}{v_o} \times \left[J_1(v_o)\cos\left(\frac{1}{2}u_o\right) + u\int_0^1 \rho^2 J_1(v_o\rho)\sin\left(\frac{1}{2}u_o\rho^2\right) d\rho\right]. \tag{3.13}$$

Using the Eq. (3.12) and integral by parts again, we get finally,

$$C(u_o, v_o) = \frac{\cos\left(\frac{1}{2}u_o\right)}{u_o} u_1(u_o, v_o) + \frac{\sin\left(\frac{1}{2}u_o\right)}{u_o} u_2(u_o, v_O), \tag{3.14}$$

where $u_n(u_o, v_o)$ is the first Lommel function, as discussed in Ref. [12] and [18],

$$u_n(u_o, v_o) = \sum_{s=0}^{\infty} (-1)^s \left(\frac{u_o}{v_o}\right)^{n+2s} J_{n+2s}(v_o). \tag{3.15}$$

Similar we get that,

$$S(u_o, v_o) = \frac{\sin\left(\frac{1}{2}u_o\right)}{u_o} u_1(u_o, v_o) + \frac{\cos\left(\frac{1}{2}u_o\right)}{u_o} u_2(u_o, v_o), \tag{3.16}$$

Substituting Eqs. (3.14-3.16) into Eq. (3.8), we get the first of our semi-analytical solution of Eq. (3.7),

$$U_1(u_o, v_o) = \frac{u_1(u_o, v_o) + u_2(u_o, v_o)i}{u_o} e^{-\frac{iu_o}{2}}, \tag{3.17}$$

Where the subscript "1" indicate that this is the first solution of the integral. Note that the term $J_{n+2s}(v_o)$ in the first Lommel function converge to zero. In the illumination area, $|u| < |v|$, the term $\left(\frac{u}{v}\right)^{n+2s}$ converge to zero, too. It therefore is appropriate to use this solution in the illumination area [19][20].

II. The second Lommel solution

We now provide a solution to the Eq. (3.7) using the second Lommel function. Integral by parts is still used. But with $A(x) = J_0(v_o\rho)$ and $B'(x) = \cos(\frac{1}{2}u_o\rho^2)$ [21], or using alternatively the equations from Ref. [12], Chap. 16, Page 537-542,

$$\frac{v_1(u_o, v_o) - v_0(u_o, v_o)i}{u_o} e^{\frac{-iu_o}{2}} = \int_1^\infty J_0(v_o\rho) e^{\frac{-iu_o\rho^2}{2}} \rho d\rho, \tag{3.18}$$

$$\int_0^\infty J_0(v_o\rho) e^{\frac{-iu_o\rho^2}{2}} \rho d\rho = -\frac{i}{u_o} e^{\frac{iv_o^2}{2u_o}}. \tag{3.19}$$

$v_n(u_o, v_o)$ is the second Lommel function,

$$v_n(u_o, v_o) = \sum_{s=0}^{\infty} (-1)^s \left(\frac{v_o}{u_o}\right)^{n+2s} J_{n+2s}(v_o). \tag{3.20}$$

Using Eqs. (3.19) - (3.18), the solution is,

$$U_2(u_o, v_o) = -\frac{i}{u_o} e^{\frac{iv_o^2}{2u_o}} - \frac{v_1(u_o, v_o) - v_0(u_o, v_o)i}{u_o} e^{\frac{-iu_o}{2}}. \tag{3.21}$$

Using the second Lommel function the term $\left(\frac{v_o}{u_o}\right)^{n+2s}$ appears instead of $\left(\frac{u_o}{v_o}\right)^{n+2s}$ in Eq. (3.15), this indicates the solution converges in the shadow area. So we propose to use this solution in the geometric shadow area[19][20].

III. Nijboer's solution with Zernike polynomials

We now turn out attention to employing the third solution. In Eq. (3.7) we have the term

$$e^{-\frac{iu_o\rho^2}{2}} = e^{-\frac{iu_o}{4}} e^{-\frac{iu_o}{4}(2\rho^2 - 1)}. \tag{3.22}$$

and from Ref. [18] and Ref. [22], we note that,

$$e^{ic(2\rho^2-1)} = \sqrt{\frac{\pi}{2c}} \sum_{n=0}^{\infty} i^n (2n+1) J_{n+\frac{1}{2}}(c) P_n(2\rho^2-1), \tag{3.23}$$

where P_n is the Legendre polynomial, which is related to the Zernike polynomial (Ref. [23]), as follows

$$P_n(2\rho^2-1) = R_{2n}^0(\rho), \tag{3.24}$$

$R_{2n}^0(\rho)$ is the Zernike radial polynomials,

$$R_n^{|m|}(\rho) = \sum_{l=0}^{\frac{n-|m|}{2}} (-1)^l \frac{(n-l)!}{l!\left(\frac{|m|+n}{2}-l\right)!\left(\frac{n-|m|}{2}-l\right)!} \rho^{n-2l}, \tag{3.25}$$

where "!" is the factorial operation. n and $(n-m)$ are non-negative integers, and $(n-m)$ must be even. Substituting Eqs. (3.22-3.24) into Eq. (3.7), and introducing the parameter $c = -u/4$,

$$U(u_o, v_o) = \sum_{n=0}^{\infty} e^{-\frac{iu_o}{4}} \sqrt{\frac{2\pi}{-u_o}} i^n (2n+1) J_{n+\frac{1}{2}}\left(-\frac{u_o}{4}\right) \int_0^1 R_{2n}^0(\rho) J_0(v\rho) \rho d\rho. \tag{3.26}$$

Another relationship, which is important in the diffraction theory of aberrations [23], is that

$$\int_0^1 R_{2n}^0(\rho) J_0(v\rho) \rho d\rho = (-1)^n \frac{J_{2n+1}(v_o)}{v_o}. \tag{3.27}$$

Substituting Eq. (3.27) into Eq. (3.26) gives the third solution for Eq. (3.7),

$$U_3(u, v) = \sum_{n=0}^{\infty} e^{-\frac{iu_o}{4}} \sqrt{\frac{2\pi}{-u_o}} (-i)^n (2n+1) J_{n+\frac{1}{2}}\left(-\frac{u_o}{4}\right) \frac{J_{2n+1}(v_o)}{v_o}. \tag{3.28}$$

IV. Cao's solution

More recently Cao defined a family of generalized Jinc functions $w_n(v)$ as follows [17],

$$w_n(v_o) = \frac{1}{v_o^{2n+2}} \int_0^{v_o} x^{2n+1} J_0(x) dx. \tag{3.29}$$

The zero order($n = 0$) of this function is the traditional Jinc function. We rearrange the integral as follows, with $\rho = x/v_o$,

$$w_n(v_o) = \int_0^1 \rho^{2n+1} J_0(v_o\rho)d\rho. \tag{3.30}$$

This can also be rewritten as the form of a polynomial [17],

$$w_n(v_o) = \sum_{m=0}^{n} (-2)^m \frac{n!}{(n-m)!} \frac{J_{m+1}(v_o)}{v_o^{m+1}}. \tag{3.31}$$

Using the Taylor series expansion for $e^{-\frac{1}{2}iu_o\rho^2}$, it can be shown that

$$U(u_o, v_o) = \sum_{n=0}^{\infty} \frac{\left(-\frac{1}{2}iu_o\right)^n}{n!} \int_0^1 \rho^{2n+1} J_0(v_o\rho)d\rho. \tag{3.32}$$

Substituting Eqs. (3.30-3.31) into Eq. (3.32) we get a fourth solution of the form,

$$U_4(u_o, v_o) = \sum_{n=0}^{\infty} \frac{\left(-\frac{1}{2}iu_o\right)^n}{n!} \sum_{m=0}^{n} (-2)^m \frac{n!}{(n-m)!} \frac{J_{m+1}(v_o)}{v_o^{m+1}}. \tag{3.33}$$

Apply these four solutions we can calculate the intensity distribution, to within a constant multiplicative factor, at any plane behind the lens,

$$I(u_o, v_o) = |U(u_o, v_o)|^2. \tag{3.34}$$

3.1.2 A comparison of some analytical solutions

It must be emphasized that, all of the four semi-analytical solutions, derived in the previous section, are mathematically equivalent. They all involve summing over an infinite number of Bessel functions to solve the integral exactly. In practice, however, only a finite number of terms can be used in the calculation and so the rate of convergence of each solution is of very real practical importance. It is shown that the rate at which each method converges has strong spatial dependence and we establish a relationship between the spatial location and the convergence rate for each solution.

In order to compare the relative performance of each of the different semi-analytical solutions we have found it useful to compare the results to those found by directly numerically integral the diffraction integral Eq. (3.7), which has been discussed in Chap. 2. The result U_N for such a calculation is presented in Fig. 3.2.

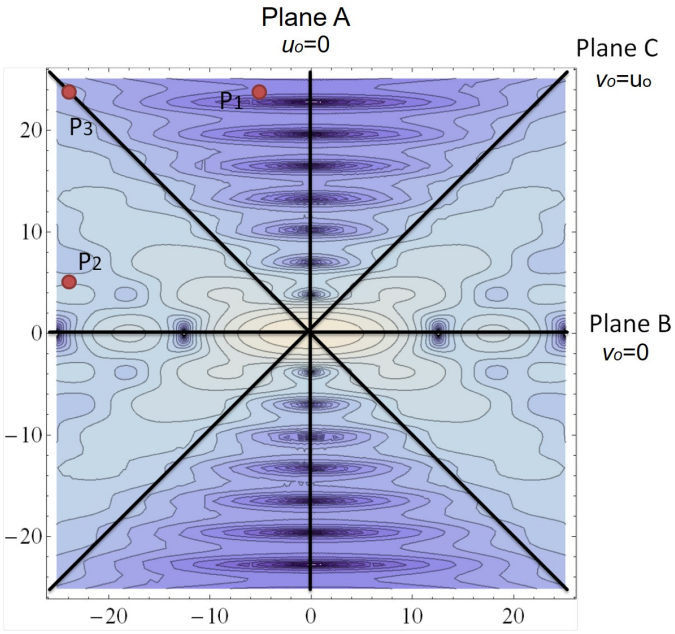


Fig. 3.2: A contour plot of $|U_N|^2$ in the focal region.

We first however compare the analytical solutions for given numbers of series terms along different cross-sections through the focus. We then examine how many series terms are required for each solution to meet the same convergence criteria at a set of specific points, P_1, P_2 and P_3, see Fig. 3.2. Various cross-sections and spatial locations are identified in Fig. 3.2. After that, we examine the differences between each analytical solution and the numerically calculated results and produce a set of error maps so that the significants of these errors can be visualized. Finally, the time taken to calculate the results using each of the analytical solutions are demonstrated.

3.1.2.1 A comparison of the semi-analytical solutions along several cross-sections through the focal region

We begin our comparison of the semi-analytical solutions by examining their predictions along three different cross-sections, each of which passes through the focal point. Three different cross-sections are shown in Fig. 3.2 and defined as: A. along the focal plane ($u_o = 0$), B. along the optical axis ($v_o = 0$) and C. along the boundary of the illumination area ($|u_o| = |v_o|$).

A. Along focal plane ($u_o = 0$)

For each solution only the first five terms of the analytical series is used for the calculation and we examine $|U(u_o, v_o)|^2$ over the range $-30 < v_o < 30$. We note that when $u_o = 0$, one must perform a limiting operation in order to arrive at the

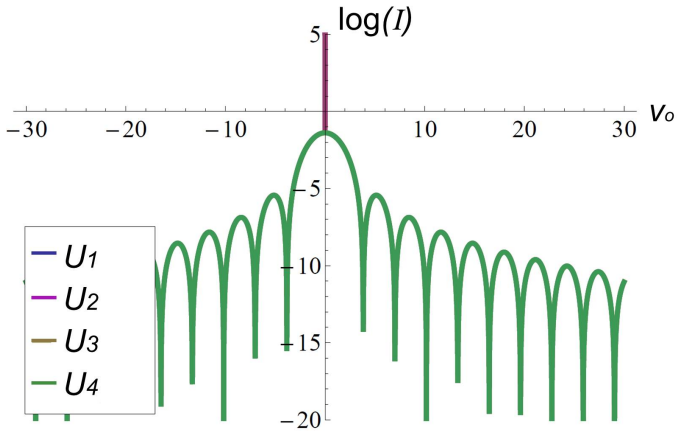


Fig. 3.3: Intensity along the focal plane ($u_o = 0$) when five series terms are used in the calculations.

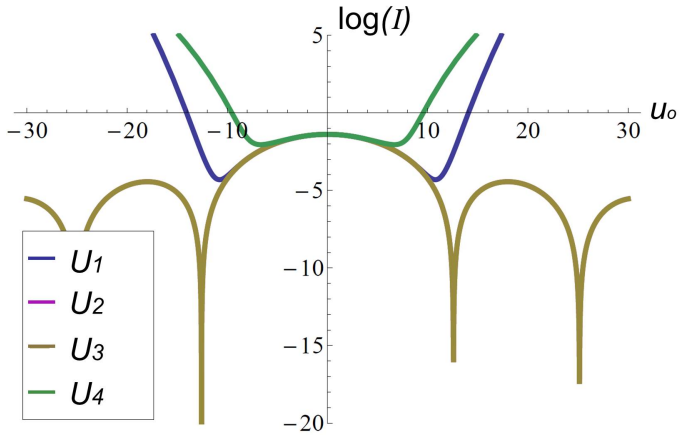


Fig. 3.4: Intensity distribution along the optical axis ($v_o = 0$) retaining five series terms.

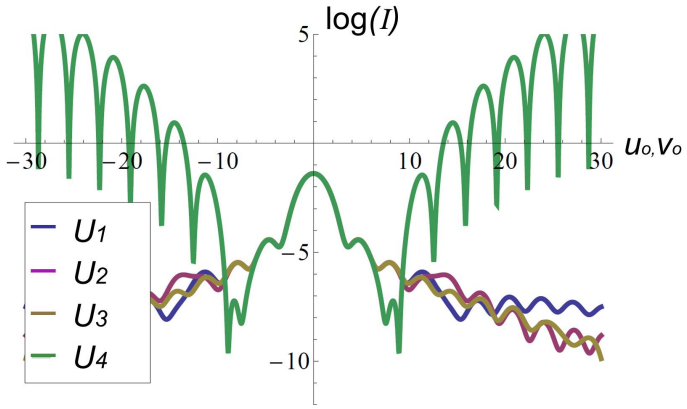


Fig. 3.5: Intensity at the boundary ($|u_o| = |v_o|$) of the illumination area. Again only five series terms are used in the calculations.

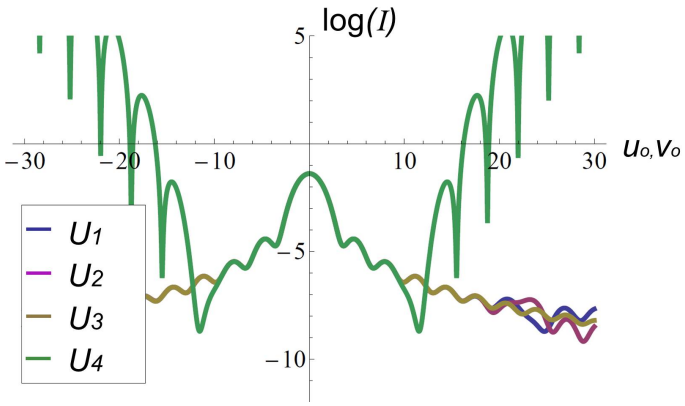


Fig. 3.6: Intensity at the boundary ($|u_o| = |v_o|$) of the illumination area with ten terms.

correct result for the first Lommel function, second Lommel function and Zernike polynomial solutions. This is because of the presence of the $1/u_o$ term. In practice we use $u_o = 0.0001$ close to $u_o = 0$. Fig. 3.3 shows the log of the intensity distribution in the focal plane. The solutions, $|U_1|^2$, $|U_3|^2$ and $|U_4|^2$ converge quickly and closely agree with each other in Fig. 3.3. However the second Lommel function, $|u_2|^2$, does not converge. As when $u_o \approx 0$, and $v_o \gg u_o$, the term $(v_o/u_o)^{n+2s}$ in Eq. (3.20) requires that more terms be included so as to achieve convergence.

B. Along the optical axis ($v_o = 0$)

As in the previous calculation along the focal plane, we again use the first five series terms in each of the analytical solutions. The calculation range is $-30 < u_o < 30$. Again due to the presence of the $1/v_o$ term, we only calculate the result for $v_o = 0.0001$ instead of $v_o = 0$. The comparison is shown in Fig. 3.4. In this region $|U_2|^2$ and $|U_3|^2$ converge quickly (brown-yellow distribution). But for the other two, five terms are not sufficient to ensure that the analytical solution converges. The U_4 solution converges over the range $-6 < v_o < 6$ when 5 terms are used while the $|U_1|^2$ solution performs better converging over the range $-9 < v_o < 9$ (see the green and blue plots in Fig. 3.4 respectively). Therefore we conclude that it is better to use $|U_2|^2$ or $|U_3|^2$ near the optical axis, i.e. when $v_o = 0$.

C. Along the boundary of the illumination area ($|v_o| = |u_o|$)

In this region we have $|v_o| = |u_o|$, and the calculation range is $-30 < v_o = u_o < 30$. The results from the four analytical solutions are presented in Fig. 3.5. The result predicted by $|U_4|^2$ converges (green plot) for spatial locations near the focal region, however does poorly for most other spatial locations along the cross-section. In fact each solution converges at different rates along this cross-section. To examine this feature in more detail, we calculate the intensity distributions again, however now using ten terms, see Fig. 3.6. It is clear that the $|U_3|^2$ distribution does not change,

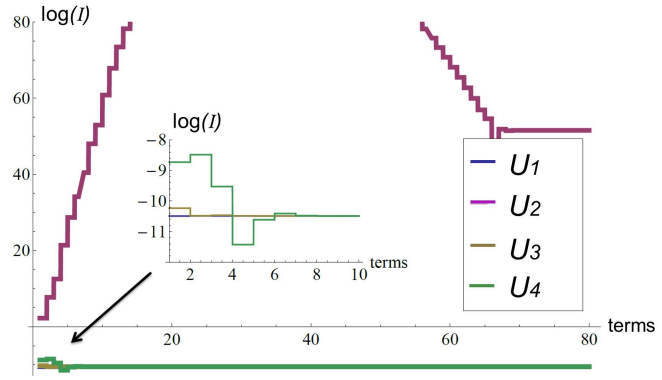


Fig. 3.7: Converge process at P_1 as a function of series terms.

however the additional terms improve the convergence of the other three solutions. So we see that the $|U_3|^2$ solution converges most quickly along this boundary region while the $|U_4|^2$ solution performs the worst in this region.

3.1.2.2 Convergence rate at three specific spatial locations: P_1, P_2 and P_3

In the previous section we examined the performance of the analytical solutions along three different cross-sections of the focal distribution all of which pass through the focus. In this section we wish to examine in more detail how each solution converges as a function of the number of series terms retaining. We choose three specific spatial locations as indicated in Fig. 3.2.

P_1. Near the focal plane $|u_o| = 5$, $|v_o| = 25$ From Fig. 3.3 we see that at the focal plane the performance of the U_1, U_3 and U_4 solutions are good, while the U_2 solution does not converge. A question arises as to whether we can make U_2 converge by simply including more series terms? In the Fig. 3.7 we calculate the intensity at point $P_1(|u_o| = 5, |v_o| = 25)$ near the focal plane, and we see that the U_1, U_3 and U_4 solutions converge rapidly. Increasing the number of terms used with the U_2 solution does not seem to help and the solution does not converge. This lack of convergence is numerical in nature and is related to the number of significant figures that can be used to represent a rational number in a computer. "The floating-point machine reals follow the IEEE standard Double Format using 53 bits of machine storage (including one hidden bit) with a machine epsilon of 2^{-52} (which is approximately 2.2×10^{-16})"[24], and hence the inaccuracy is of the order 10^{-16}. In most cases this difference does not cause numerical instability, however here, if $v_o \gg u_o$, the second Lommel solution produces values that are larger than 10^{16}. In the U_2 series solution at P_1, successive series terms are very large and opposite in sign. Adding large numbers of opposite sign means that the resulting small round-off errors can quickly lead to numerical stability problems. While there are means of improving

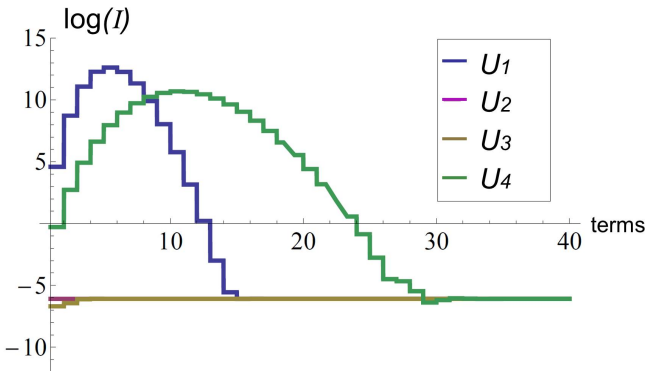


Fig. 3.8: Converge process at P_2 as a function of series terms.

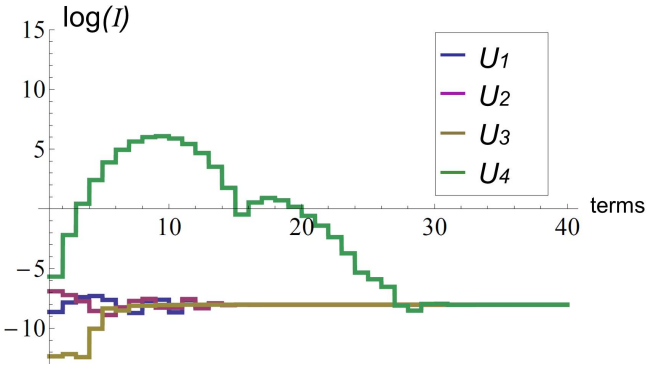


Fig. 3.9: Converge process at P_3 as a function of series terms.

the numerical accuracy, they are not pursued here and we say that second Lommel function does not provide a numerically stable solution in this region.

P_2. Near the optical axis $|u_o| = 25$, $|v_o| = 5$

Now we turn to look at the point near the optical axis $|u_o| = 25$, $|v_o| = 5$. From Fig. 3.8 we see that both the U_2 and U_3 solutions converge quickly needing only a few series terms to reach a final stable value. The U_1 and U_4 solutions converge more slowly. If we compare the manner in which the U_2 solution in Fig. 3.7, and the U_1 and U_4 solutions in Fig. 3.8, converge, we can see a similar trend. Although the U_1 and U_4 solutions converge for all the cases presented here, we expect from this observation that similarly there are regions where these solutions will fail due to similar numerical stability reasons. We have not found a simple means of defining these unstable regions. Hence we adopt the rule that if the solution does not stabilize as the series are increased, one should choose an alternative solution.

P_3. At the boundary of geometric shadow $|u_o| = 25$, $|v_o| = 25$

This point is located away from the optical axis and the focal plane, and we can see that more terms are required to achieve the same level of accuracy (see Fig. 3.9). The U_1 and U_2 solutions need about fourteen terms to reach the correct answer. The U_3

solution converges the fastest, needing only ten terms. The U_4 solution needs more than thirty terms to converge. We note the nature of the convergence is similar in form to U_2 in Fig. 3.7. As we move out further along curve C, and away from the focus, this U_4 solution becomes increasingly unstable and will eventually reach a point where it dose not converge due to the digital limit discussed before.

3.1.2.3 Error maps

In this section we determine regions where the solutions converge well and where the maximum errors are to be found. As reference we use the directly numerically integrating of Eq. (3.7) in Chap. 2. We refer to this numerical solution as U_N, 10000 sampling points of ρ are taken to ensure that the replicas are far away from each other enough, so that an accuracy of 10^{-6} is guaranteed [25].

We compare the four analytical solutions, each calculated using only ten terms, with a numerical solution U_N. In Figs. 3.10 - 3.13 the resulting errors can be seen and regions of convergence identified. We make the following observations:

1. The $U_1(u, v)$ solution converges well within the illuminated region behind the lens.
2. The opposite appears to hold for the $U_2(u, v)$ solution, where the largest errors are within the illuminated region. It does however converge well in the geometrical shadow.
3. The Nijboer solution provides the most stable solution over the ranges we have examined. The error increases as one moves away from the focus of the lens.
4. The Cao solution seems to be have a performance that is a combination of the Lommel solutions. It is not as robust as the Nijboer and the errors appear to be symmetrical about the $u = 0$ plane.

From these results we conclude that while all the solutions are mathematically equivalent they require a finite number of series terms to be used in practical calculations. When only a finite number of terms are used, the Nijboer solution has a very robust convergence properties in the whole focal region.

3.1.2.4 A comparison of the calculation time

In addition to the accuracy, the total time taken to calculate a numerical result from the analytical solutions is an important parameter. Here we examine how the computation time varies for each of the four analytical solutions. We calculate 100×100 points distributed in the focal region $-20 < u < 20, -20 < v < 20$. The

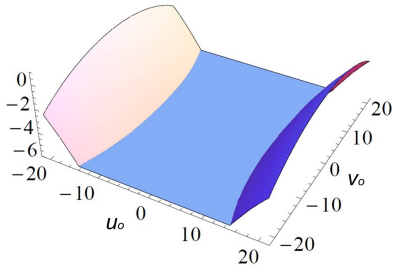


Fig. 3.10: Error map of solution U_1, $\log(|U_1 - U_N|)$ is plotted.

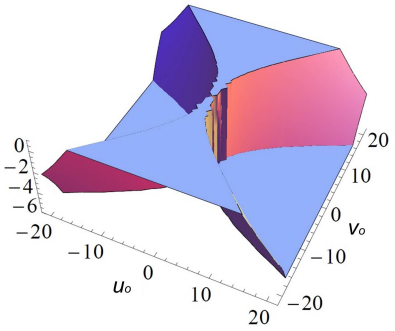


Fig. 3.11: Error map of solution U_2, $\log(|U_2 - U_N|)$ is plotted.

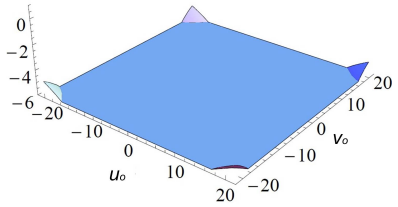


Fig. 3.12: Error map of solution U_3, $\log(|U_3 - U_N|)$ is plotted.

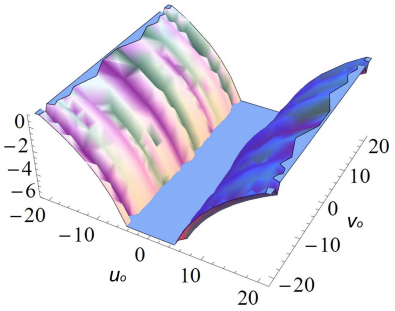


Fig. 3.13: Error map of solution U_4, $\log(|U_4 - U_N|)$ is plotted.

computer we used is an Intel(R) Core(TM) i7-2600K, and the computing platform is Matlab. The computation time for the four solutions are plotted as a function of the number of series terms, see Fig. 3.14. The U_4 solution takes the most time of all, then follow the Nijboer's solution and two Lommel solutions, the time spent of the three solutions are similar. The Nijboer's solution performs very well, not only the fast computation time but also the relative stable convergence properties.

3.2 A focal region calculation based on the Extended Nijboer Zernike theory (ENZ)

In the section above we discussed the semi-analytical solutions of a perfect converging lens, although all four solutions mathematical equivalent are, the convergence properties of them are totally different between each other, and we found that the Nijboer's solution with Zernike polynomials is the most robust solution of all. In this section, we introduce the focal region calculation of aberrated lens based on ENZ theory.

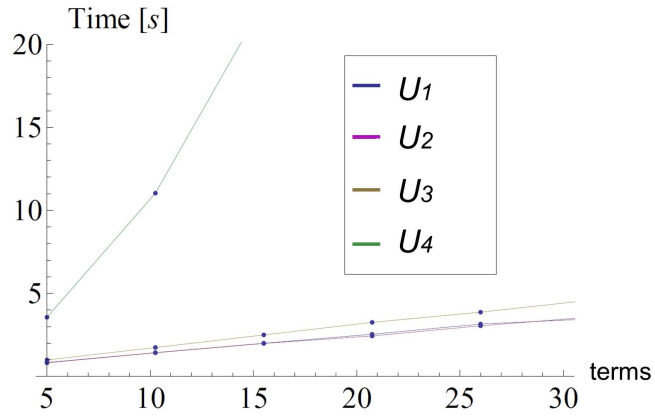


Fig. 3.14: Comparison of the computation time for the four semi-analytical solutions.

3.2.1 The Extended Nijboer Zernike theory

In 1942 Nijboer established the diffraction theory of abberation using Zernike polynomials [23]. He gave the analytical solution of the optical point spread functions(PSF) at the focus plane in an aberrated system. In 2002 Janssen extended Nijboer's work to include the defocus case[26, 27, 28]. Later, S. van Haver and J.J.M. Braat used ENZ to solve different kinds of diffraction integral for systems with high numerical aperture(NA) and vector fields[29, 30, 31, 32], and introduced truncation technique to accelerate the calculation of the algorithm[33]. Compare to traditional numerical techniques, ENZ theory as a semi-analytical technique of the diffraction integral, has a very high accuracy of the results. The calculated diffraction field is continuous, and suffers no replicas effect.

The Zernike polynomials are a sequence of polynomials that are orthogonal on the unit disk and are named after optical physicist Frits Zernike[34]. The polynomials are defined as following,

$$\begin{cases} Z_n^m(\rho,\phi) = R_n^{|m|}(\rho)\cos(m\phi), & m \geq 0 \\ Z_n^m(\rho,\phi) = R_n^{|m|}(\rho)\sin(m\phi), & m < 0 \end{cases} \tag{3.35}$$

where $R_n^{|m|}(\rho)$ is the Zernike radial polynomials, which was defined in Eq. (3.25). Note that different orders of Zernike polynomials are orthogonal to each other and correspond to the different lens aberrations[35](see Fig. 3.15). The ENZ can solve different kinds of diffraction integral in the focal region. The basic ideal is to decompose different terms in the integral into Zernike series, then using the following equation[36, 23],

$$\int_0^{2\pi} \exp(iv_o\rho\cos(\theta-\phi))\cos(m\phi)d\phi = 2\pi i^m J_m(v_o\rho)\cos(m\theta), \tag{3.36}$$

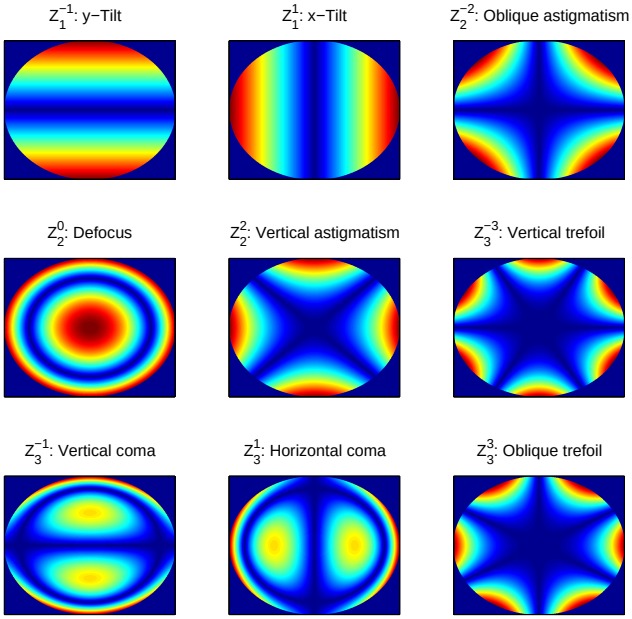


Fig. 3.15: The 2-10th order of Zernike polynomials Z_n^m, corresponding to the lens aberrations.

$$\int_0^{2\pi} \exp(iv_o\rho\cos(\theta-\phi))\sin(m\phi)d\phi = 2\pi i^m J_m\left(v_o\rho\right)\sin(m\theta), \tag{3.37}$$

$$\int_0^1 R_n^m(\rho)J_m(v_o\rho)\rho d\rho = (-1)^{\frac{n-m}{2}}\frac{J_{n+1}(v_o)}{v_o}, \tag{3.38}$$

to get the analytical express.

Assuming we have a focus systems, the diffraction integral for this system can be written as [32],

$$U(v_o,\theta) = \int_0^1\int_0^{2\pi} A_f(\rho)F(\rho)P(\rho,\phi)\exp\left[iv_o\rho\cos(\theta-\phi)\right]\rho d\rho d\phi. \tag{3.39}$$

Where $A_f(\rho)$ and $F(\rho)$ are the algebraic factor and focal factor. $P(\rho,\phi)$ is the aberration term. If we have a system without aberration, $P(\rho,\phi)=1$, Otherwise, the aberration can be written in terms of Zernike polynomials,

$$P(\rho,\phi) = \sum_{n,m}\beta_n^m Z_n^m(\rho,\phi), \tag{3.40}$$

which β_n^m is the complex Zernike coefficients. The algebraic factor (related to radiometric effect) and focal factor (related to defocus) are different based on the diffraction approaches we use. For example, in the Debye approximation with a small defocus, these two factor are [32],

$$A_f(\rho) = 1, \tag{3.41}$$

$$F(\rho) = \exp(id_f\rho^2); \tag{3.42}$$

in case of a large defocus, the two factor are changed to,

$$A_f(\rho) = \frac{1}{\sqrt{1-\mathrm{NA}^2\rho^2}}, \tag{3.43}$$

$$F(\rho) = \exp\left[i\frac{d_f}{1-\sqrt{1-\mathrm{NA}^2}}\left(1-\sqrt{1-\mathrm{NA}^2\rho^2}\right)\right]. \tag{3.44}$$

where d_f is the defocus term,

$$d_f = k(z-f)\left(1-\sqrt{1-\mathrm{NA}^2}\right), \tag{3.45}$$

and NA$= f/2a$ is the object space numerical apertures. Using the mathematical treatment in [32], and the Clebsch-Gordan coeffcient in [37], the product of Zernike radial polynomials, algebraic factor $A(\rho)$ and focal factor $F(\rho)$ is written as,

$$A_f(\rho)F(\rho)R_n^{|m|} = \sum_{h,l} A_{2l,n,h}^{0mm} w_t R_h^{|m|}(\rho), \tag{3.46}$$

$A_{2l,n,h}$ is the Clebsch-Gordan coeffcient and w_t is the sum of different Clebsch-Gordan coeffcients related to $A_f(\rho)F(\rho)$, see Eq. (35-57) in [32] for detail.

Substitute Eq. (3.35,3.40,3.46) into Eq. (3.39), and use the analytical solution of the integral in Eq. (3.36-3.38), the semi-analytical solution of Eq. (3.39) can be calculated.

The calculation of ENZ can be divided into three steps: Decomposing the aberration term $P(\rho,\phi)$ into finite orders of Zernike polynomials; Using Clebsch-Gordan coeffcient to combined the product of two or more Zernike radial polynomials to one Zernike cirle polynomias; Accumulating the analytical results of integrals of every orders of Zernike polynomials. ENZ theory provides a semi-analytical treatment to calculate the focal region for different systems, not only the low-NA systems with Rayleigh integral, but also the high-NA systems with vector fields, of cause with more

calculation efforts. Compare to traditional numerical calculation, ENZ has a very high accuracy of 10^{-6} in appropriate location ($|d_f| < 25$) [32], and has no replicas in the whole output region.

3.2.2 Simplified Extended Nijboer Zernike theory (SENZ)

ENZ offers an effective and accurate semi-analytical solution of different diffraction integrals. However the implementation of the resulting algorithm is complex, and needs a triple summation of a series of Zernike radial polynomials. In this section, we propose a simplified ENZ technique, which greatly reduces the complexity of the ENZ.

In traditional ENZ, we use a Zernike decomposition to fit the aberration term $P(\rho, \phi)$. With the help of the Clebsch-Gordan coeffcient we combine the aberration term with the algebraic factor $A_f(\rho)$ and the focal factor $F(\rho)$. At the end we get the analytical 3D distribution at the focal region of the system. In our approach, we decompose the product of $P(\rho, \phi)$, $A_f(\rho)$ and $F(\rho)$ instead of Eq. (3.40),

$$P(\rho, \phi)A_f(\rho)F(\rho) = \sum_{n,m} \beta_n^m Z_n^m(\rho, \phi). \tag{3.47}$$

Then substituting Eq. (3.47) into Eq. (3.39),

$$U(v_o, \theta) = \int_0^1 \int_0^{2\pi} \sum_{n,m} \beta_n^m Z_n^m(\rho, \phi) \exp\left[iv_o\rho\cos(\theta - \phi)\right] \rho d\rho d\phi, \tag{3.48}$$

$$\begin{aligned} U(v_o, \theta) = &\int_0^1 \int_0^{2\pi} \sum_{n,m\geq 0} \beta_n^m R_n^{|m|}(\rho)\cos(m\phi) \exp\left[iv_o\rho\cos(\theta - \phi)\right] \rho d\rho d\phi + \\ &\int_0^1 \int_0^{2\pi} \sum_{n,m<0} \beta_n^m R_n^{|m|}(\rho)\sin(m\phi) \exp\left[iv_o\rho\cos(\theta - \phi)\right] \rho d\rho d\phi. \end{aligned} \tag{3.49}$$

With the help of Eq. (3.36-3.37), we analytically integrate the angle $d\phi$,

$$\begin{aligned} U(v_o, \theta) = &\int_0^1 \sum_{n,m\geq 0} \beta_n^m R_n^{|m|}(\rho) 2\pi i^m J_m\left(v_o\rho\right)\cos(m\theta)\rho d\rho + \\ &\int_0^1 \sum_{n,m<0} \beta_n^m R_n^{|m|}(\rho) 2\pi i^m J_m\left(v_o\rho\right)\sin(m\theta)\rho d\rho. \end{aligned} \tag{3.50}$$

At the end, the product of Zernike radial polynomials $R_n^{|m|}$ and Bessel function J_m is integrated analytically according to Eq. (3.38),

$$U(v_o,\theta) = 2\pi i^m (-1)^{\frac{n-m}{2}} \frac{J_{n+1}(v_o)}{v_o} \left(\sum_{n,m\geq 0} \beta_n^m \cos(m\theta) + \sum_{n,m<0} \beta_n^m \sin(m\theta) \right). \tag{3.51}$$

In this simplified approach, the calculation process is divided into two steps: Decomposing the term $P(\rho,\phi)A_f(\rho)F(\rho)$ into finite orders of Zernike polynomials; Accumulating the analytical results of integrals of every orders of Zernike polynomials. Without the Clebsch-Gordan coeffcient step, the programming complexity and the calculation time are greatly reduced.

3.2.3 Comparison of the ENZ and the SENZ approaches

As shown in the two section above, both the ENZ and the SENZ are the semi-analytical approaches. By projecting different field into a Nijboer-Zernike basis set, then using the known analytical solution of the Zernike radial polynomials integration, the diffraction field at the output plane is calculated semi-analytically. In this section we compare these two approaches in two points of view: the computation speed and the accuracy.

The total computation time t_t consists of two parts, t_d by decomposing the term ($P(\rho,\phi)$ in ENZ or $P(\rho,\phi)A_f(\rho)F(\rho)$ in SENZ), and t_a by accumulating different orders of Zernike radial polynomials. In the decomposing part, we use Shakibaei's recursive technique to generate different orders of R_n^m [38, 39], the computational complexity is significant reduced and this computation time can be omitted. By performing the least mean squared in Matlab, we get the complex-valued Zernike coefficients β_n^m. The computation time depends on two variables: on the one hand, the number of points of the to be decomposed term $P(\rho,\phi)$ or $P(\rho,\phi)A_f(\rho)F(\rho)$; on the other hand, the maximal orders of the Zernike radial polynomials. If the expression to be decomposed is similar to a lens abberation, as in Fig. 3.15, only a few Zernike terms are required to fit the field well. However if we want to decompose a field like a speckle field, that has random amplitudes and phases[40, 41, 42], a large number of Zernike terms must be taken, which leads to a significant increase of the computation time. In our comparison, if the size of the decomposed term and the maximal orders of the Zernike radial polynomials are the same, we have $t_d^{ENZ} = t_d^{SENZ}$.

The second part of the computation time is the accumulation part. If we compare the computation time vertically, it is proportional to the number of points which

need to be calculated in the output plane. If we compare the time horizontally, the ENZ has a triple summation of the Zernike radial polynomials while the SENZ has only a single summation. Therefore within the same order of Zernike polynomials and the same number of the points in the output plane, $t_a^{ENZ} > t_a^{SENZ}$. Therefore if we calculate a diffraction plane of a system, SENZ is faster than the ENZ. However if the focal region of a system need to be calculated, which means the diffraction field with for example L planes with different distances, the situation is reversed. In ENZ the defocus term is contained in step 2 and 3, to calculate the output field at different distances, the decomposing step only need to be done once, so the total computation time is written as,

$$t_t^{ENZ} = t_d^{ENZ} + L \times t_a^{ENZ}. \tag{3.52}$$

However in SENZ, the defocus term is combined with the aberration term, the decomposing step need be performed for every distance, and the total computation time is written as,

$$t_t^{SENZ} = L\left(t_d^{SENZ} + t_a^{SENZ}\right). \tag{3.53}$$

Although $t_a^{ENZ} > t_a^{SENZ}$, if we want to calculate L output planes, ENZ will be a more efficient way to take instead of SENZ.

As semi-analytical solutions instead of closed form solutions, both approaches have systematical error in the calculation. Note in the algorithms, the accumulating step is a purely analytical step (step 3 in ENZ and step 2 in SENZ), there is no numerical error. Although step 2 in ENZ use finite polynomials to approach the product of two Zernike radial polynomials, there is a guideline to chose the required terms to keep the error to an arbitrary small amount [33]. The only step which could introduce large error is the decomposing step. By comparing the accuracy of these two approaches, we just need to test how well the Zernike polynomials fits the terms $P(\rho,\phi)A_f(\rho)F(\rho)$ or $P(\rho,\phi)$.

Assuming we have a system with the focal length $f = 20$ mm, wave length $\lambda = 600$ nm, NA=0.6, and a small abberation term $P(\rho,\phi)$(Fig. 3.16). By performing the Debye approximation we calculate the diffraction plane at a distance $z = 19.995$ mm using Eq. (3.43) and Eq. (3.44). With ENZ approach we use the maximum Zernike radial polynomials $n_{max} = 20$ to fit $P(\rho,\phi)$, and the reconstructed field is shown in Fig. 3.17. From the figures can be seen that the Zernike fit works very well, the maximum difference between the aberration term and the reconstructed aberration term is of the order of 10^{-5}. For the SENZ approach, we use the same maximal Zernike radial polynomials $n_{max} = 20$ to fit $P(\rho,\phi)A_f(\rho)F(\rho)$ instead. The

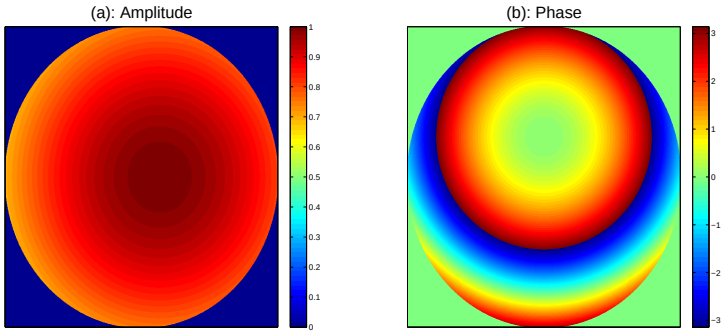


Fig. 3.16: The amplitude and phase of the aberration term $P(\rho, \phi)$.

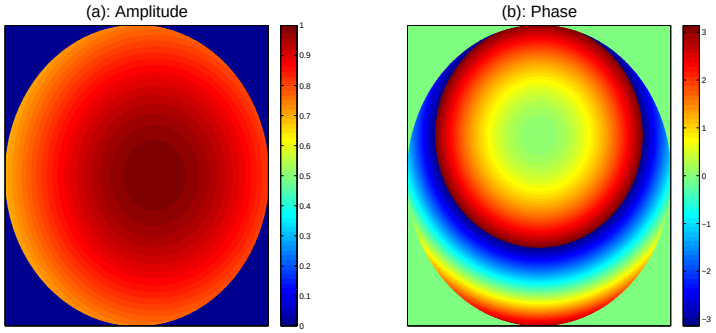


Fig. 3.17: The reconstructed amplitude and phase of the aberration term $P(\rho, \phi)$ with a Zernike fit of $n_{max} = 20$.

amplitude and the phase of the term $P(\rho, \phi)A_f(\rho)F(\rho)$ is shown in Fig. 3.18, and its reconstruction is plotted in Fig. 3.19. The decomposing process of SENZ works as well as ENZ, the maximal difference between the original term and the reconstructed term is also of the same order of 10^{-5}. We also chose some other examples with different kinds of aberration term $P(\rho, \phi)$, and found that the reconstruction error of both ENZ and SENZ methods are similar. From the analysis above we conclude the accuracy of both techniques are at the same level. In Fig. 3.20, the intensity distribution at a distance of $z = 19.995$ mm is plotted, the results of ENZ and SENZ coincide very well, the difference of them varies between 10^{-5} and 10^{-4}.

In this section, we compare ENZ and SENZ in two different ways. For the accuracy, both approaches perform very well at the same level. In the computation time point of view, if only a single diffraction plane is calculated, SENZ is faster. If we want to calculate the whole diffraction field with different diffraction distances z, then the ENZ is more efficient than the SENZ.

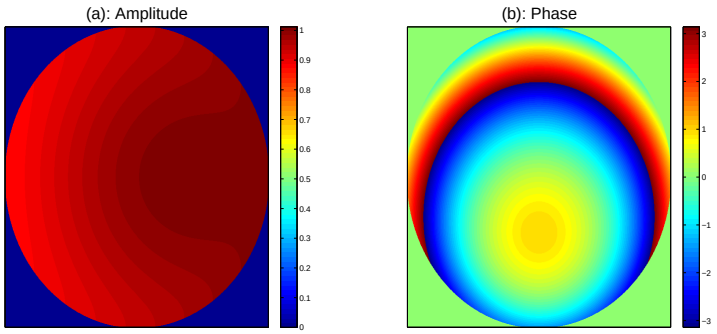


Fig. 3.18: The amplitude and phase of the term $P(\rho,\phi)A_f(\rho)F(\rho)$.

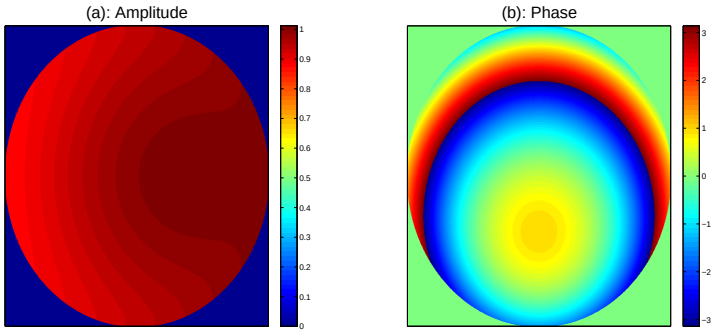


Fig. 3.19: The reconstructed amplitude and phase of the term $P(\rho,\phi)A_f(\rho)F(\rho)$ with a Zernike fit of $n_{max} = 20$.

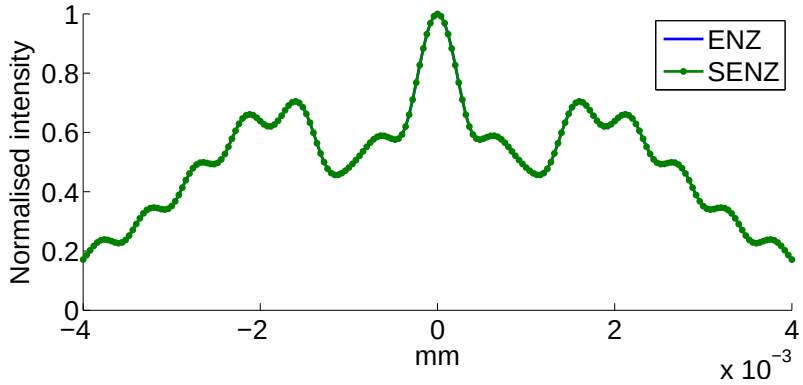


Fig. 3.20: A comparison of the intensity distribution ($Y = 0$, -4×10^{-3} mm$< X < 4\times10^{-3}$ mm is plotted) using ENZ and SENZ. System variables: focal length 20 mm, NA=0.6, defocus 0.005 mm, wavelength 600 nm, aberration term see Fig. 3.17.

3.3 The Fresnel transform as a projection onto a series of Zernike polynomials

3.3.1 Mathematical derivation

In the previous section we have discussed the ENZ and SENZ methods for calculating the optical distribution in the focal region of a lens system. In this section we try to extend the scope of the ENZ based technique and calculate the output of a general function using the Fresnel transform. In our approach, we use the Fresnel transform in cylindrical coordinate system from Eq. (2.15), and use Eq. (3.4) to normalise the radius r,

$$U(R,\theta) = a^2 e^{\frac{ik}{2z}R^2} \int_0^1 \int_0^{2\pi} u(\rho,\phi) e^{\frac{ik}{2z}(a\rho)^2} e^{\frac{-ikRa\rho}{z}\cos(\phi-\theta)} \rho d\rho d\phi. \tag{3.54}$$

Note in this approach, the input field is defined in a circular aperture, the optical field outside the aperture is set to zero. The idea here is similar to SENZ, we combine the term $u(\rho,\phi)e^{\frac{ik}{2z}(a\rho)^2}$ and project it onto a series of Zernike radial polynomials[43],

$$\bar{u}(\rho,\phi) = u(\rho,\phi)e^{\frac{ik}{2z}(a\rho)^2} = \sum_{m\geq 0,n}^{\infty} \beta_n^m R_n^{|m|}(\rho)\cos(m\phi) + \sum_{m<0,n}^{\infty} \beta_n^m R_n^{|m|}(\rho)\sin(m\phi). \tag{3.55}$$

Substituting Eq. (3.55) into Eq. (3.54) then integrating over ϕ and ρ using Eq. (3.36-3.38), we get the following result,

$$U(R,\theta) = 2\pi a^2 e^{\frac{ik}{2z}R^2} i^m (-1)^{\frac{n-m}{2}} \frac{z J_{n+1}(-kRa/z)}{-kRa} \times \left(\sum_{m\geq 0,n}^{\infty} \beta_n^m \cos(m\theta) + \sum_{m<0,n}^{\infty} \beta_n^m \sin(m\theta) \right). \tag{3.56}$$

Theoretically we need to sum over an infinite number of n, m terms to generalize the pupil function $\bar{u}(\rho,\phi)$. In practice only a finite number of terms are taken to describe the field with $\bar{u}_f(\rho,\phi)$. The average differences between the fitted field and the original field is defined as,

$$E_f = \left\langle |\bar{u}(\rho,\phi) - \bar{u}_f(\rho,\phi)| \right\rangle, \tag{3.57}$$

where "$\langle\rangle$" is the ensemble average operation. As demonstrated in the last section, E_f plays an important role in determining the accuracy of the technique. Similar to the SENZ method, the number of necessary terms of Zernike radial polynomials is dependent on the complexity of the pupil function. In order to verify our technique,

Table 3.1: Fitting error and maximal relative error between our ENZ method and the reference second Lommel result using maximal Zernike radial polynomials $n_{max} = 15$

input field	fitting error E_f	maximal relative error E
200×200 pixels	1.0%	3.5%
400×400 pixels	0.96%	1.7%
600×600 pixels	0.93%	1.1%
800×800 pixels	0.92%	0.7%

we use the maximal relative error E to quantify the accuracy of our technique,

$$E = \max\left[\frac{|U_e(R,\theta) - U_r(R,\theta)|}{U_r(R,\theta)}\right]. \tag{3.58}$$

which $U_e(R,\theta)$ is calculated from Eq. (3.56), and $U_r(R,\theta)$ is the reference solution.

3.3.2 Simulation result

As mentioned before, there is generally no analytical solution for the Fresnel transform, so we chose a test case with a semi-analytical result. We consider a converging wave front with a circular aperture of radius 2 mm, which is focused at a distance of $f = 400$ mm. Therefore we have a continuous input field $u(\rho,\phi) = e^{-ik\rho^2/2f}$. A strongly defocused output field is calculated at the distance $z = 600$ mm. As a reference solution we use the second Lommel function Eq. (3.21) which converges quickly in this area. In order to get the complex Zernike coefficient β_n^m we need to sample the input field and fit it with Zernike radial polynomials. We use the fitting error E_f to estimate how well we describe the aperture function. At first, we sample the input field with a matrix of 200×200 pixels, the fitting error is $E_f = 1\%$ using $n_{max} = 15$ Zernike radial polynomials, and the maximal relative error to the reference solution is $E = 3.5\%$, see Tab. 3.1. Then we sample the input field more finely with 400×400, 600×600, 800×800 pixels, the fitting error takes a nearly constant value of 0.9%. However E is reduced to 1.7%, 1.1% and 0.7%. For the continuous input field, the finer we sample the original field $\bar{u}(\rho,\phi)$, more accurate is the fitted field $\bar{u}_f(\rho,\phi)$ to the original field. Therefore in this case, the sampling rate of the input field has a significant impact on the accuracy.

In the second test case we investigate the propagation of a sampled input field. The second test case corresponds to the common situation that the input field is represented by discrete values. Fig. 3.21 shows the amplitude and the phase of an input field, its amplitude is a shifted Gaussian distribution, and the phase is given as a sine wave along the x and y axis. Here we calculate the Fresnel transform at a distance of $z = 600$ mm using our technique. In this case, there is no analytical solution, and it is difficult to find a perfect reference. Therefore we compare the result

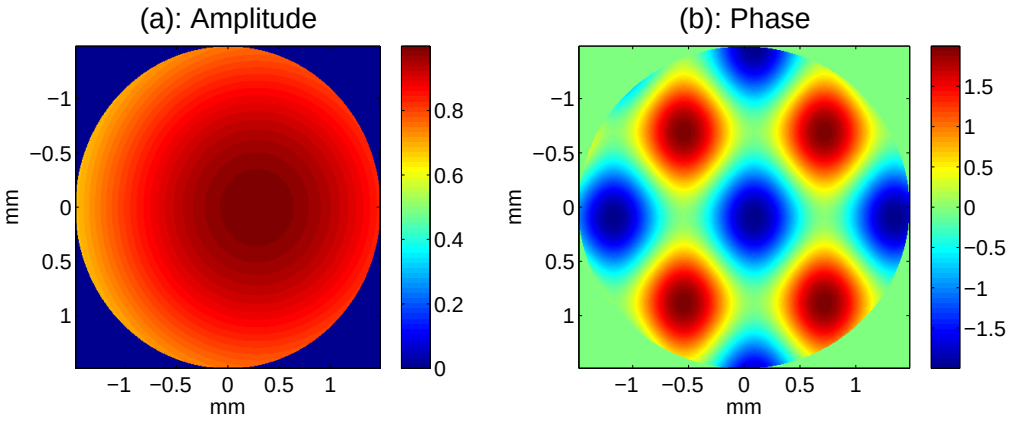


Fig. 3.21: Amplitude and phase distribution of the input field $u(\rho, \phi)$.

using ENZ technique to the result using point by point integration as shown in Eq. (2.2) with $\Delta x = \Delta y = 7.4\mu m$ and each axis is sampled with 400 samples. The results are presented in Fig. 3.22 (a) and are nearly identical. However the point by point numerical integration has a low accuracy, and can not be used as a perfect reference. From Eq. (3.55) we see that the accuracy of our method increases with the number of Zernike terms. We thus fit the input field with $n_{max} = 40$, and take this as our reference. In this case the fitting error is reduced to approximately 0.1%. Fig. 3.22 (b) show the fitting error and the maximal relative error to the reference for $n_{max} = 15$, 20, 25, 30, 35. With more Zernike polynomials, the fitting error reduces significantly, from 60% to 0.3%, also the maximal relative error reduces to 0.15% for $n_{max} = 35$. It shows that our technique works for converging but also for non-converging fields, and the results have an accuracy of 10^{-3} if the fitting error is smaller than 1%. This simulation has also coincided with the argument we made in Sec. 3.2.3, the fitting error is the only error source of our technique.

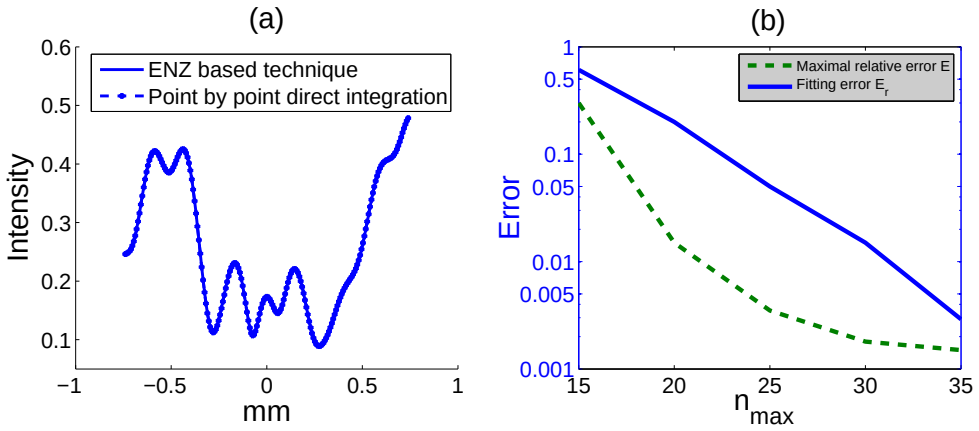


Fig. 3.22: (a): Intensity distribution across the optical axis. The curve 1 shows the results using ENZ based technique ($n_{max} = 40$), the curve 2 shows the results using point by point integration. (b): The fitting error E_f and the maximal relative error E to the reference in terms of the number of used Zernike terms.

A traditional CCD array gives a rectangular input plane. If we want to calculate the Fresnel transform using ENZ based technique, we need set the pixel values located outside the circular aperture, to 0. Also, a traditional CCD array includes digital noise. This noise can be significantly reduced by the Zernike fit. In Fig. 3.23 we use the Fresnel transform to reconstruct a hologram of a converging lens (It will be more detailed in the next chapter). The phase modulation of the lens is calculated by (a): our ENZ based technique with $n_{max} = 40$ and (b): FFT based spectral method. Because of the digital noise, the reconstruction in Fig. 3.23 (b) is not as clear as the reconstruction in Fig. 3.23 (a), whose digital noise is reduced by the fitting process.

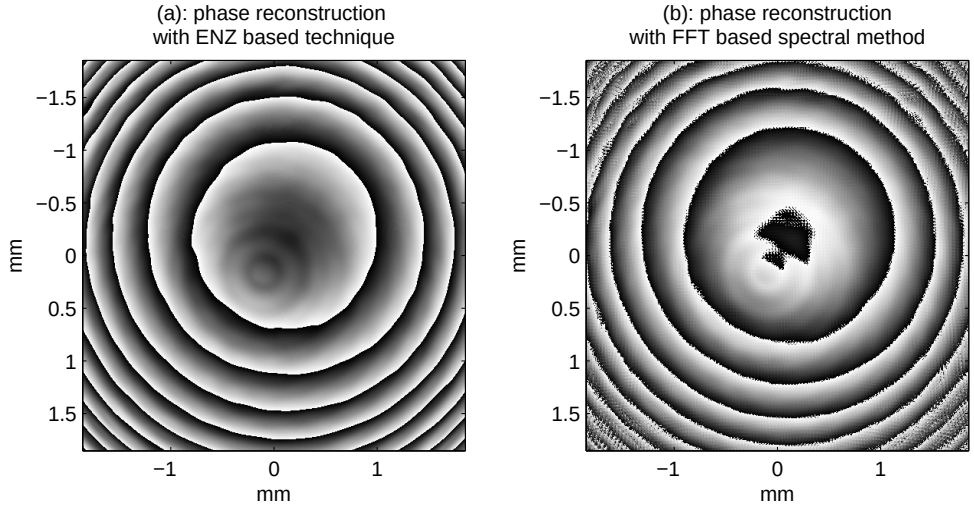


Fig. 3.23: Phase reconstruction of a hologram using (a): ENZ based technique with $n_{max} = 60$, (b): FFT based spectral method

However if we want to reconstruct a hologram of a USAF chart, it is not suitable to use our technique. In Fig. 3.24 we show the reconstuction using (a): our ENZ based technique with $n_{max} = 60$ and (b): the FFT based spectral method. It is obvious that our technique failed in this case, more terms of Zernike polynomials need to be taken to fit the term $\bar{u}(\rho, \phi)$. To obtain accurate results, a much larger number of Zernike terms would be required. The Nijboer-Zernike basis set is defined to correspond to different orders of abberation in a lens. If we want to project a field, which is similar to a lens abberation as in Fig. 3.23, to the Nijboer-Zernike basis set, only a few Zernike terms are enough to fit the field well. However if we want to project a field like a hologram of the USAF chart in Fig. 3.24, a large number of Zernike terms must be taken, which is not suitable anymore to use this technique.

In this section we have introduced a technique to calculate the optical Fresnel transform semi-analytically. Unlike the usual FFT based techniques, we can calculate the output field at arbitrary points. In our ENZ technique, the accuracy is dependant

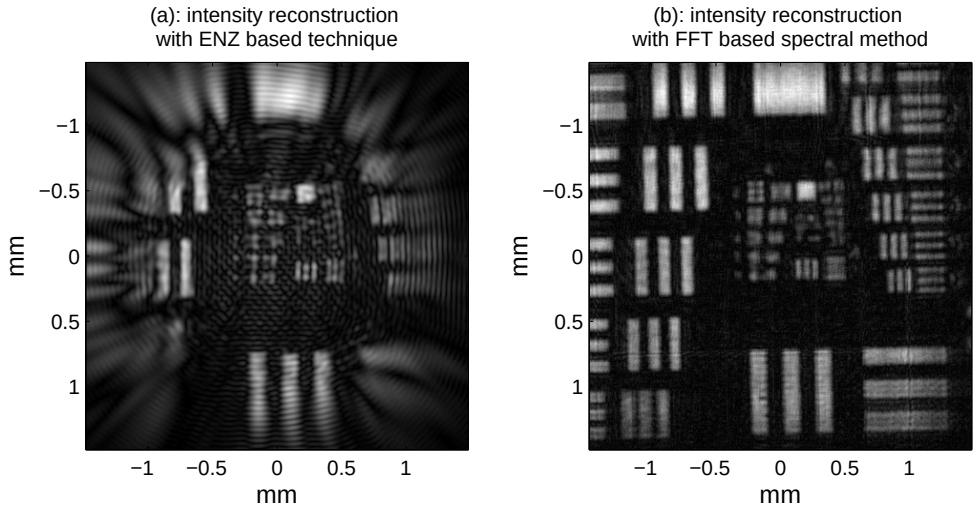


Fig. 3.24: Intensity reconstruction of USAF chart using (a): ENZ based technique with $n_{max} = 60$, (b): FFT based spectral method

on the sampling rate of the input plane and the number of Zernike terms used to fit the input plane. Therefore for some complex input field, to achieve a high level of accuracy, we need to fit the input field with more Zernike terms, which increases calculation time. In our technique, if we have enough Zernike terms to describe the pupil function $\bar{u}(\rho, \phi)$, we don't need to worry about the problems of conventional numerical methods, like replicas, limited sampling rates and the extent of the output plane. Our method has also some limitations. It works well for long distance propagation with a small circular pupil, $ka^2/z < 20$, and some field similar to the lens abberation, where only a few Zernike terms are required to fit the quadratical phase term $e^{\frac{ika^2\rho^2}{2z}}$. For the near field propagation or some complex field like a speckle field, this technique is less suitable, as a high number of Zernike terms are required for accurately fitting the input plane and it takes a long calculation time.

3.4 Conclusion

In this chapter, we have discussed the semi-analytical solution of the diffraction integrals. We begin our analysis with a typical diffraction problem, intensity distribution of focal region of a perfect converging system. By comparing the four different solutions, we found the Nijboer's solution with Zernike polynomials converges very fast at the whole focal region, performances the best. Then we investigate the ENZ theory and propose a simplified version of the ENZ. The SENZ is easy to program and has the same accuracy as the ENZ. Using the same technique as SENZ, we proposed a semi-analytical solution of the Fresnel transform.

4 Holographic imaging system

4.1 Introduction to holography

In Chap. 2 and Chap. 3, we discussed the propagation of light in wave optics using the Fresnel transform. If we know both the intensity and the phase of the optical field in a particular plane, then we can use the Fresnel transform to calculate the optical field in different planes. Here we examine how we can use this feature to build a lensless imaging system. In practice, the intensity can be easily measured by film or photoelectric sensors, however the phase is more difficult to recover. In this chapter, we will introduce and use holography to capture the phase information of the optical field, and reconstruct the object in a lens-less holographic system.

In a holographic setup, the light source is divided by a beam splitter into two parts: the object arm and the reference arm, see Fig. 4.1. The reference wave $A_r(x, y)$ is a plane wave and propagates through the optical system, until it comes to the second beam splitter. Meanwhile the object arm (also a plane wave) is incident upon the object in Fig. 4.1. We refer to the field immediately after the object as the scattered object wave $U_o(x, y)$. The scattered object wave propagates to the CCD plane after the Fresnel transform, which is written as $A_o(x, y)$. These two beams interfere with each other and the resulting intensity pattern is recorded by the electronic detection. This is a stable intensity pattern, which we call a hologram. For the traditional holography, the hologram is captured by film, and the object wave can be reconstructed by the laser illumination [44]. If we use a digital device like a CCD instead of film to capture the hologram and then reconstruct the object wave numerically, the technique is then called digital holography.

In digital holography, we have the hologram $H(x, y)$,

$$\begin{aligned} H(x,y) &= |A_o(x,y) + A_r(x,y)|^2 \\ &= |A_o(x,y)|^2 + |A_r(x,y)|^2 + A_o^*(x,y)A_r(x,y) + A_o(x,y)A_r^*(x,y) \\ &= |A_o(x,y)|^2 + |A_r(x,y)|^2 + 2A_o(x,y)A_r(x,y)\cos(\psi(x,y)), \end{aligned} \quad (4.1)$$

where A_r is the complex reference wave, A_o is the complex object wave at the sensor plane, $\psi(x, y)$ is the phase difference between the object and reference wave, and

the superscript "*" is the conjugate operation. There are three different parts in the hologram: the DC term $|A_o(x,y)|^2 + |A_r(x,y)|^2$, real image term $A_o(x,y)A_r^*(x,y)$ and the twin image term $A_o^*(x,y)A_r(x,y)$. The task of the reconstruction step in digital holography is to separate the real image term from the other terms. Once the real image term is separated, we have both the phase and amplitude of the object wave. Then we can use the Fresnel transform to propagate the object wave from the sensor plane back to the object plane, which builds a lens-less imaging system.

In this chapter, we use the phase shifting technique [45] to separate the object wave from the others (in which we take 4 captures, and for every capture the phase difference $\psi(x,y)$ is shifted over $\pi/2$), and reconstruct the object in the Fresnel domain. The resolution limits in this lens-less imaging system are discussed and experimentally approached. At the end of the chapter, we use this lens-less holographic system to measure the phase modulation of the object, and therefore indirectly measure the optical thickness of the object.

4.2 Phase Shifting Interferometry

It is important to be able to quantify, both theoretically and experimentally, the performance of coherent digital systems so that their suitability for a given metrology application can be assessed. Here a free-space inline digital holographic system is investigated. To isolate the scattered object field, Phase Shifting Interferometry (PSI) techniques are used. A typical setup is depicted in Fig. 4.1. In our setup light from a He-Ne laser is spatially filtered with a microscope objective and a pinhole (diameter of 20 microns) which cleans up the beam and improves the spatial coherence. The light emerging from the pinhole is modeled as an ideal point source which is collimated with an achromatic lens to form an approximate plane wave traveling along the optical axis. The first beam-splitter (BS1) diverts the plane wave into the object and reference arms respectively. In the reference arm the mirror M can be shifted precisely using a piezomotor to realize relative phase shifting between the reference and object arms. In the object arm, the plane wave illuminating the object, is scattered and propagates in free space until it is incident on the CCD detector (IMPERX IGV-B2020 with 2056×2060 pixels, and the pixel size is 7.4 μm). The second beam splitter is used to recombine the reference and object beams, where they interfere producing a hologram that is recorded by the digital CCD device. This intensity is in fact a digital hologram. To isolate the desired real image term we use a 4-step Phase Shifting Interferometry (PSI) technique [45] and a piezo motor (P-611.ZS) supplied by "Physik Instrumente".

A theoretical examination of the performance of such a holographic system [4, 46, 47]

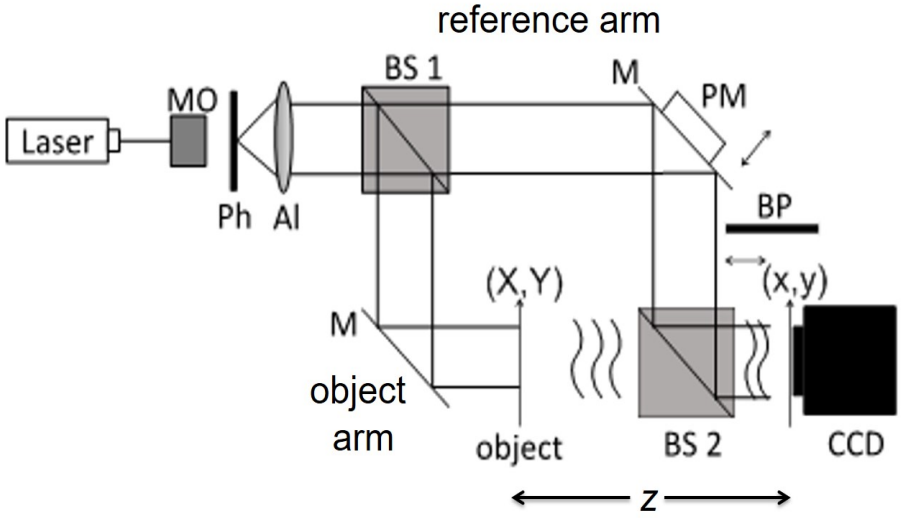


Fig. 4.1: Schematic depicting the measurement setup: M, mirror; BS, beamsplitter; MO, microscope objective; Al, Achromatic lens; Ph, pinhole; PM, piezo motor; BP, black plate (to block the reference wave, will be used for error metric to estimate the performance of PSI).

in the paraxial regime supposes that there are no measurement errors in the recording process and that the following assumptions hold:

(i) A perfectly flat reference field,

(ii) With a noise free digital capture, where no quantization (rounding off due to finite number of bits) occurs in the digital sensor, [uniform pixel characteristics are also assumed]

(iii) vibration free environment,

(iv) and a perfectly calibrated PSI technique.

If this is the case then it has been shown [48, 49, 50, 51, 52, 4, 47] that the ultimate performance of the system is by several fundamental factors:

(1) the finite extent of the CCD/CMOS array, which limits the maximum spatial frequency to approximately $D/\lambda z$, which D is the half width of the CCD/CMOS array, λ is the wavelength and z is the distance between the object and the detection;

(2) averaging of the light intensity over the finite size of the each pixel, which limits the resolution spot with the spatial frequency $1/\gamma$, where γ is the width of the active area of a pixel. This is effectively a convolution operation and will remove spatial frequencies, see [46, 47];

Table 4.1: Meaning of parameter in Eq. (4.2 - 4.5, 4.10)

$A_r(x,y)$	complex reference wave
$A_o(x,y)$	complex object wave at the camera plane
$\psi(x,y)$	initial phase difference between the object and reference wave
$\Delta\psi_o(x,y)$	phase modulation from the object
ψ_n^s	phase shift by the piezo motor
$\psi_v(t)$	phase change due axial vibration
N_n	sensor noise
τ	exposure time
$I_o^m(x,y)$	measured intensity when reference wave is blocked
$I_o^c(x,y)$	intensity as calculated from PSI capture

(3) the sampling interval or distance between the centers of each pixel, Δ, acts to limit the field of view in the reconstruction domain, to the range: $\lambda z/\Delta$.

We begin our analysis with the following holographic equation [53],

$$H_n(x,y;t) = \tau \cdot H_{DC}(x,y) + H_K(x,y) \cdot H_{AC}(x,y;t) + N_n, \tag{4.2}$$

$$H_{DC}(x,y) = |A_r(x,y)|^2 + |A_o(x,y)|^2, \tag{4.3}$$

$$H_K(x,y) = 2|A_r(x,y)||A_o(x,y)|, \tag{4.4}$$

$$H_{AC}(x,y;t) = \int_t^{t+\tau} \cos[\psi(x,y) + \Delta\psi_o(x,y) + \psi_n^s + \psi_v(t')]dt', \tag{4.5}$$

where A_r is the complex reference wave, A_o is the complex object wave at the camera plane, $\psi(x,y)$ is the systematic phase difference between the object and reference wave, $\Delta\psi_o(x,y)$ is phase of the object field that we want to measure, ψ_n^s is phase shift that we deliberately introduce with the piezo motor, $\psi_v(t)$ is a phase error introduced by axial vibration and N_n is electronic sensor noise, and τ is the sensor exposure time. We use the parameter n refers to the index each of the four holograms (see Tab. 4.1).

Under the ideal conditions outlined earlier, i.e. when the assumptions (i)-(iv) are all fulfilled, then $N_n = 0$, $\psi_v = 0$, the phase shift is precisely $\psi_n^s = n\pi/2$, and we have the ideal inline reference and object plane waves, the initial difference is constant all over the camera, we can set $\psi(x,y) = const.$ Using the four-step-phase shifting algorithm [45], we can recover the object phase distribution $\Delta\psi_o(x,y)$ and

the object amplitude $A_o(x, y)$ with the following equation,

$$A_o(x, y) = H_1(x, y) - H_3(x, y) + i\left[H_4(x, y) - H_2(x, y)\right]. \tag{4.6}$$

So the recovered intensity $I_o^c(x, y)$ and the phase $\Delta\phi_o(x, y)$ are,

$$I_o^c(x, y) = \left[H_1(x, y) - H_3(x, y)\right]^2 + \left[H_4(x, y) - H_2(x, y)\right]^2, \tag{4.7}$$

$$\Delta\psi_o(x, y) = \arg(A_o). \tag{4.8}$$

In the paraxial regime we use the Fresnel transform to recover complex object wave in the object plane with $U_o(X, Y)$ [10],

$$U_o(X, Y) = \frac{\exp(i2\pi/\lambda)}{i\lambda z} \int_{-\infty}^{\infty} A_o(x, y) \exp\left\{\frac{i\pi}{\lambda z}\left[(x - X)^2 + (y - Y)^2\right]\right\} dxdy, \tag{4.9}$$

where the uppercase (X, Y) and lowercase (x, y) refer to spatial coordinates in the object and detection (CCD) planes respectively, see Fig. (4.1). We note that the complex object field at the sensor plane, $A_o(x, y)$, is a continuous complex function that is then subject to several filtering and sampling operations before a discrete set of values are returned by the CCD device [46, 47]. The parameters associated with the filtering and sampling operations are determined by the properties of the digital recording device. We assume that these operations are constant and hence examine how the field $A_o(x, y)$ changes due to conditions (i) - (iv), see in particular [49, 51, 48].

From Eq. (4.7) we see that it is possible to separate the DC and twin image terms from the real image term and hence we can calculate the intensity $I_o^c(x, y)$ of the object wave. It is also possible to directly measure the intensity $I_o^m(x, y)$ of the object wave by blocking the reference wave. If we can recover the object wave perfectly then directly measured intensity should be identical to that recovered from the PSI process, otherwise measurement errors lead to differences between the measured intensity and recovered intensity. We can directly estimate this with the following error metric E,

$$E = \sum_{x,y} \left[I_o^c(x, y) - I_o^m(x, y)\right]^2. \tag{4.10}$$

This error metric, E, is a direct measure of any experimentally errors and system noise introduced by the holographic recording process.

In the following section, we discuss in detail each of the idealized conditions: (i) - (iv) in the following sections individually. We provide a theoretical analysis of

the errors introduced to the holographic measurement by each non-ideal condition and provide experimental results to support the description.

4.3 Vibration

In this section we examine the effect that vibration has on our system, which has also been discussed in these references[54, 55, 56, 57, 58, 59]. We assume that any vibration has a small pure amplitude with lower than micrometer. Also we assume that the lateral influence of the vibrations can be neglected as far they are much smaller than the pixel. This implies that the vibration will not influence the H_{DC} and H_K terms in Eq. (4.2), and only the axial vibration would affect the interference term H_{AC} with $\psi_v(t)$.

We now write this vibration term in a Fourier series expansion [56],

$$\phi_v(t) = \sum_{m=0}^{\infty} b_m \sin(m2\pi vt). \tag{4.11}$$

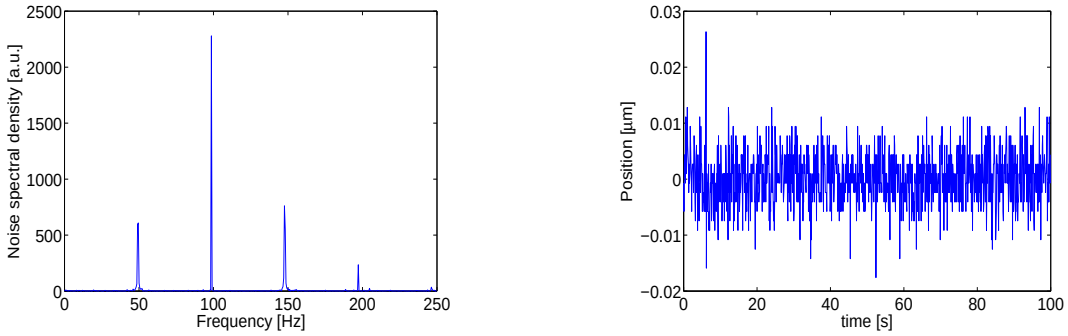
The digital sensor integral such signal over the exposure time τ, as shown in Eq. (4.5). It works as a low pass filter with the bandwidth $1/\tau$ [60]. Considering only one Fourier component, we make, $V = mv$ and $b = b_m$, and ignoring all the other noise. Then we consider only one pixel and rewrite the hologram intensity $H_n(x, y; t)$ as $H_n(t)$,

$$H_n(t) = H_K \int_t^{t+\tau} \cos\left[\Delta\psi_o + b\sin(2\pi Vt')\right] dt' + \tau \cdot H_{DC}. \tag{4.12}$$

As stated earlier, low amplitude vibrations are considered and we have negligible lateral effect of the vibration. Thus the terms H_{DC} and H_K are assumed to be constant. As such, the vibration only effect the interference term $\cos\left[\Delta\psi_o + b\sin(2\pi Vt)\right]$. Using a standard trigonometric relation, we can rewrite Eq. (4.12) as,

$$H_n(t) = H_K \int_t^{t+\tau} \{\cos\Delta\psi_o \cos\left[b\sin(2\pi Vt')\right] - \sin\Delta\psi_o \sin\left[b\sin(2\pi Vt')\right]\} dt' + \tau \cdot H_{DC}. \tag{4.13}$$

Now using a Bessel function expansion of the time-dependent trigonometric functions, see Chap. 2, Page 23 of Ref. [12], the terms inside the square brackets of Eq. (4.13)



(a) (b)

Fig. 4.2: (a) Shows the spectral density function of the vibration noise of 50 Hz by measuring the fluctuation of one pixel intensity, then do the Fourier transform of the intensity series; (b) shows the axial position fluctuation of piezo motor in still.

can be expressed as,

$$\cos\left[b\sin(2\pi Vt')\right] = J_0(b) + 2\sum_{n=1}^{\infty} J_{2n}(b)\cos(2\pi\cdot 2nVt'), \tag{4.14}$$

$$\sin\left[b\sin(2\pi Vt')\right] = 2\sum_{n=1}^{\infty} J_{2n-1}(b)\cos\left[2\pi(2n-1)Vt'\right], \tag{4.15}$$

where $J_n(b)$ is the Bessel function with n-th order. Making use of these relationships, the hologram intensity with first three harmonic terms is given by,

$$H_n(t) = \tau\cdot H_{DC} + H_K \int_t^{t+\tau} [\cos\Delta\psi_o J_0(b) - 2\sin\Delta\psi_o J_1(b)\cos(2\pi Vt')+$$
$$2\cos\Delta\psi_o J_2(b)\cos(2\pi\cdot 2Vt') - 2\sin\Delta\psi_o J_3(b)\cos(2\pi\cdot 3Vt') + ...]dt'. \tag{4.16}$$

With this equation, we analyze the vibration for different cases.

4.3.1 Low frequency vibration

Now, let us examine the case of low frequency vibration ($V < 1/\tau$) such as that of an optical table. For low frequency we approximately have,

$$\int_t^{t+\tau} \cos(2\pi Vt')dt' \approx \tau\cos(2\pi Vt), \tag{4.17}$$

then Eq. (4.16) is rewritten as following,

$$H_n(t) = \tau \cdot H_{DC} + \tau \cdot H_K \cdot [\cos\Delta\psi_o J_0(b) - 2\sin\Delta\psi_o J_1(b)\cos(2\pi V t) + 2\cos\Delta\psi_o J_2(b)\cos(2\pi \cdot 2Vt) - 2\sin\Delta\psi_o J_3(b)\cos(2\pi \cdot 3Vt) + ...]. \quad (4.18)$$

From the equation above we see that, if we have a vibration source with the frequency V, then we expect vibration of the output signal with frequencies, $V, 2V, 3V...$ and with amplitudes $H_K \sin\Delta\psi_o J_1(b)$, $H_K \cos\Delta\psi_o J_2(b)$, $H_K \sin\Delta\psi_o J_3(b)$....

We tested the predictions of Eq. (4.18) using a vibrating mechanical shaker (in our case a loud-speaker), oscillating with a sinusoidal tone of 50 Hz. Then we capture the intensity value of a single pixel every 1 ms for 2 seconds and do the Fourier transform to get the spectral density function of the signal. As Fig. 4.2 (a) shows, in addition to the original vibration frequency V, there are doubled and tripled frequency components $2V, 3V$ and so on. The amplitudes of these harmonic frequencies depend on the Bessel function and the original object phase.

4.3.2 High frequency vibration

For the high frequency vibration ($V > 1/\tau$) of optical table, Eq. (4.16) can be simplified since the harmonic terms are filtered by the low pass filter (temporal integrating process of the sensor), we have

$$H_n = \tau[H_{DC} + H_K \cdot J_0(b)\cos\Delta\psi_o]. \quad (4.19)$$

The intensity is now time independent, however the contrast of the interference is multiplied by $J_0(b)$ and hence is determined by the amplitude, b, of the vibration.

4.3.3 Piezo-motor jitter (low amplitude vibration)

In the experiment we use a piezo motor P-611.ZS operating under a closed loop control system from "Physik Instrumente" to implement phase stepping. To perform phase shifting the movement of the mirror must be repeatedly controlled and the step size must less than that of a wavelength. Although piezo motors are highly accurate, there is still a motor jitter effect. Leaving the motor rest in a fixed position, we read the sequential position value returned by the closed loop system over 100 s. In Fig. 4.2 (b) we see the random fluctuation. The standard deviation of the fluctuation from our piezo motor is about 5 nm. We can considerate this fluctuation as a kind of vibration. A cycle phase shift 2π correspond the movement of the piezo motor of $4d = 448$ nm, d is the piezo motor step size (as will be discussed

later in Section 4.3.2). By the uncertainty of 5 nm (standard deviation of the fluctuation) of our piezo motor, we have the movement amplitude $b = 2\pi \times 5$ nm / 448 nm=0.07. Substituting $b = 0.07$ into Eq. (4.18-4.19), we see that for such small b, the zero order Bessel function $J_0(b) \approx 1$ and other order Bessel function $J_n(b) \approx 0, n = 1, 2, 3...$, which results that all the harmonic terms are approximately 0, and the system has not been much changed by the fluctuation of the piezo motor.

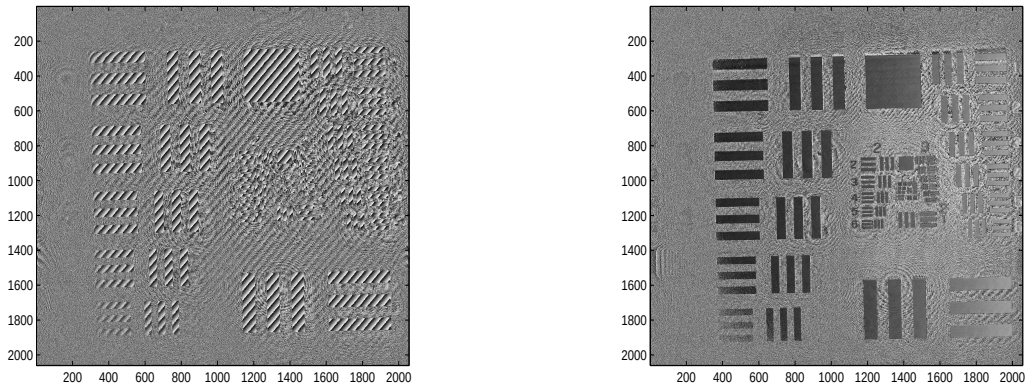
In this section, a vibration sources and their effects were discussed and analysed. The resulting mathematical model describing these sources was also presented. Using this model, motor jitter was found not to influence the imaging system. Remotely capture of holograms was found to aide reduction of ambient vibration sources.

4.4 Systematic phase errors

4.4.1 Phase error in the illuminating beam

In an in-line PSI system, a flat reference plane wave is of particular important. Ideally the illuminating beam should also be a plane wave and both should be propagating exactly along the optical axis. In this case, the phase difference $\psi(x, y)$ is constant, so we can get the clear phase modulation of the object. In our setup, presented in Fig. 4.1, a number of optical elements are used to produce a plane wave. These include a microscope objective, a pinhole and an achromatic lens, and the beams must also pass through two beam-splitters. In addition to this, there inevitably exists a small angle between the reference and objects wavefields. Such errors in the illumination wavefields result in a non-uniform initial phase difference over the sensor's active area. The resulting reconstructed phase contains the non-uniform initial phase difference between the object and reference wavefields, $\psi(x, y) + \Delta\psi_o(x, y)$.

In order to get the desired $\Delta\psi_o(x, y)$, we must know the initial phase difference $\psi(x, y)$ to compensate the recovered phase. Zernike fitting has been used to compensate the imperfect plane wave to get the phase modulation of $\Delta\psi_o(x, y)$ [61, 62]. Here we use another convenient compensation technique. In this technique the object is removed from the object wavefield and a hologram is recorded using the PSI technique described here. Employing Eq. (4.6,4.8), the initial phase, $\psi(x, y)$ is calculated. Having determined the initial phase, the object is placed into the object wavefield and a PSI hologram is captured. In our case, the object is a USAF chart. We run the PSI, and compensate for $\psi(x, y)$ from the recovered phase. With this compensated phase information, substituted into Eq. (4.9), we get the reconstructed USAF chart. Fig. 4.3 (a) shows the phase distribution without the compensation of



(a) (b)

Fig. 4.3: Phase reconstructions before (a) and after (b) the compensation. There are many fringes in (a) and no fringes in (b). The numbers on x and y axis presents the pixel number.

the initial phase difference, there are many fringe at the object bars, which gives the wrong phase modulation $\Delta\psi_o(x, y)$ from the object. Fig. 4.3 (b) presents the phase distribution with the compensation, we get the clear phase modulation of the object.

Using the compensation method described here, the object modulation is measured and can be done so using a nonuniform beam provided the maximal spatial frequency of the illumination wavefield is lower than the reconstructible frequency of the holographic system, see Points (1)-(3) in Section 4.1. This is limited by the diffraction and sensor pixel size as mentioned in Section 4.1. A clear example where this is not suitable is using an illumination wavefield containing very high spatial frequencies, such as a speckle illumination. Similarly, if we have a very good illuminating beam, we can also compensate any reference wave, whose maximal spatial frequency is lower than the reconstructible frequency.

4.4.2 PSI phase stepping error

As described above we use a piezo motor to move the mirror in the reference wave path. The laser we used is a HeNe laser with 633 nm wave length, which means we need change the reference optical wave path with a step size ϕ_n^s of $\lambda/4 = 158.25$ nm. Due to the geometrical setup, we need move the piezo motor $d = \lambda/(4\sqrt{2}) = 112$ nm every step theoretically (See Fig. 4.1, piezo motor (PM) moves with d, the optical path length of the reference wave changes with $\sqrt{2}d$). However it is difficult to move exactly this distance, and errors so introduced are referred to as phase stepping errors.

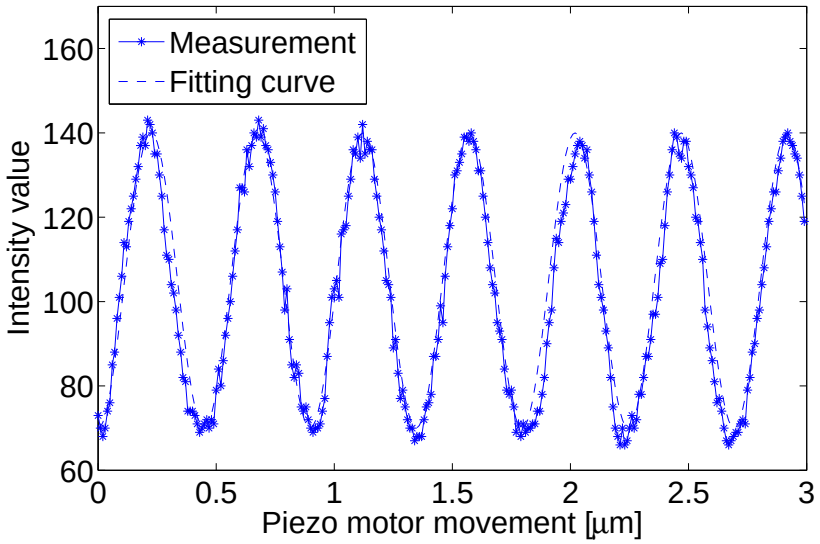


Fig. 4.4: Pixel intensity (0-255) change during the phase shift: Solid line with star is the measurement curve, and the dash line is the fitting curve.

In the setup we also need calibrate the piezo motor, to ensure that the optical path difference between every step is $\lambda/4$. From Eq. (4.5) we see that, the intensity of the hologram should take a sinusoidal appearance as we increase the phase difference by moving the piezo motor. From calculating the period, we can get the right distance of movement for the piezo motor.

In the calibrating we move every steps 10 nm for 300 steps, the intensity change of one pixel is shown in Fig. 4.4, as we expected, they are sinusoidal. By fitting this curve with a sinus curve, using least mean square method, we get a period of about 0.45 μm, corresponding the phase shift of 2π. Therefore by every step we need shift the phase with $\pi/2$, and we should move 0.45 μm/4=0.1224 μm which fits our theoretical distance $d = 0.122$ μm very well. We thus minimize the PSI error and now neglect it as a noise source in our system.

4.5 Detector noise

4.5.1 Statistical properties of detector noise

Another important factor which affects our reconstruction is the CCD/CMOS digital noise [63]. In modern optical systems discrete digital devices for measuring intensity distributions like CCD/CMOS array are widely used. The intensity incident on the CCD/CMOS array is averaged over the spatial extent of each pixel for a given exposure time. During image capture, the electronic detection noise appears. This

detection noise is random variation of brightness in images, it contains many different noise sources, such as white noise, salt-and-pepper noise, pink noise and in particular shot noise. The influence of different noise sources depends on several factors such as the exposure time, gain, pixel size, temperature etc. We expect that the intensity value measured by each pixel will fluctuate over time because of such noise. Some methods can be used post-capture to reduce noise, like resizing the original image [64] and using a median kernel filter [65], however these methods tend to lower the achievable resolution limit of the holographic imaging system.

In order to reduce the noise effect, first of all, we need to analyze the property of the noise. As shown in Fig. 4.1, we block the reference wave with the black plate. Then we use camera to capture only the fluctuation of object intensity by taking N sequential frames $I_n(x, y)$, where n refers to the frame index, from 1 to N. In order to quantify the noise level, we define the fluctuation as the relative standard deviation (RSD) of the measured intensity $I_n(x, y)$, and the RSD is

$$RSD(I_n(x, y)) = \frac{SD[I_n(x, y)]}{\langle I_n(x, y)\rangle}, \tag{4.20}$$

which $SD(\bullet)$ is the standard deviation. According to measuring the RSD of the pixel, we get the noise level of our signal.

In our previous work[66], we found the RSD is only depending on the intensity of the measured pixel, the brighter pixels have the lower RSD of the fluctuation, the RSD seems to be primarily dependent on the number of photons each pixel gets. In order to decrease the RSD we can increase the pixel value by having a longer exposure time or equivalently increasing the power of the light source. The higher the photon count the lower RSD of the signal, properties that are reminiscent of shot noise. The standard deviation (SD) of shot noise is proportional to the square root of intensity,

$$SD = c\sqrt{I}, \tag{4.21}$$

which c is one constant. So the RSD of the shot noise is,

$$RSD = \frac{c\sqrt{I}}{I} = \frac{c}{\sqrt{I}}. \tag{4.22}$$

To examine this further we examine the RSD noise level for a range of different intensities. In Fig. 4.5 we can see the two curves fitting well with each other, which indicates that the main source of our noise is shot noise. In order to decrease the RSD, we know from the analysis above, the intensity value should be kept in a high

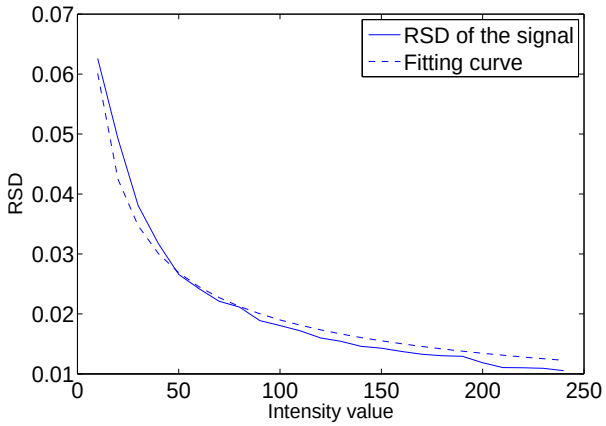


Fig. 4.5: Solid curve: RSD of pixels of different intensity value. Dashed curve: Fitting curve $c/\sqrt{I}$.

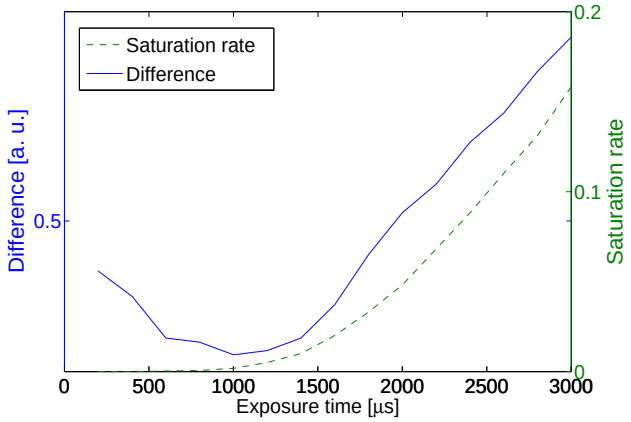


Fig. 4.6: Solid line: Intensity difference between the measurement and the recovered one at different exposure time. Dash line: Saturation percent of the image.

level, either increasing the laser power or increasing the exposure time. However we can not increase the intensity endlessly, it is limited by saturation.

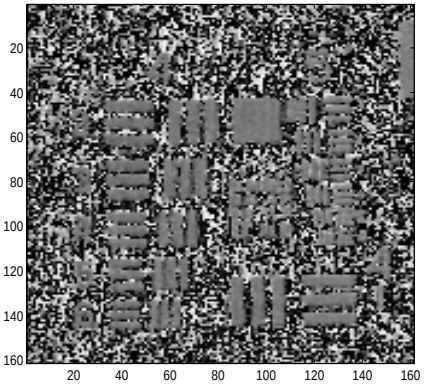
4.5.2 High Dynamic Range Holography (HDRH)

From Section 4.4.1 we know that the dominating electronic detection noise is shot noise, where the signal noise rate (SNR) of the shot noise is proportional to the square of the intensity measurement value. In order to increase the SNR, we can either increase the laser power or exposure time. In our experiment we adjust the sensor exposure time to change the intensity value from low to high, and perform the PSI measurement, to determine, which exposure time gives the best reconstruction.

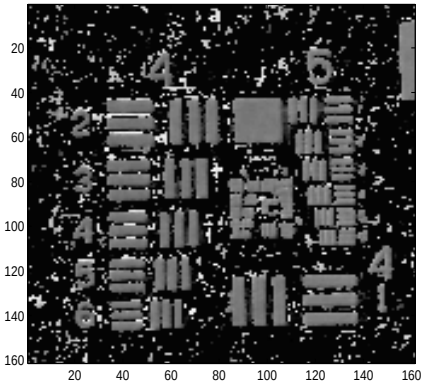
As discussed in Section 4.1 Eq. (4.10), we can contrast the recovered intensity I_o^c using Eq. (4.7) with the measured intensity I_o^m by blocking the reference wave

to estimate the error in the system. It must be emphasized, that both I_o^c and I_o^m suffer from the detector noise. In our experiment we measure the object intensity I_o^m several times by blocking the reference wave, then take the average of these intensity patterns. We say in this case, the detector noise of the measured object intensity I_o^m has been reduced due to the average process, and can be treated as the reference. The exposure time is changed from 200 μs to 3000 μs, we contrast the recovered intensity and the measured intensity to evaluate the quality of the reconstruction. The differences E, in Eq. (4.10), are shown in Fig. 4.6 with the solid line. As the exposure time increases, the difference between the two decreases. The result of this is a better reconstruction, because of the higher SNR of the shot noise. A turning point is encountered on the curve in Fig. 4.6. At an exposure time of 1000 μs, the difference between 1000 μs and 3000 μs rises dues to the saturation of the sensor pixels. Sensors have a base two dynamic range with intensity values falling between a minimum (usually 0) and a maximum. In our case, the maximum value is $(2^8 - 1) = 255$. By increasing the exposure time, some pixel values go up to 255, they are saturated, so the difference begins to rise. This effect can be seen as a dash line in Fig. 4.6, the turning point of the solid line, is the same point where the pixel start to become saturated.

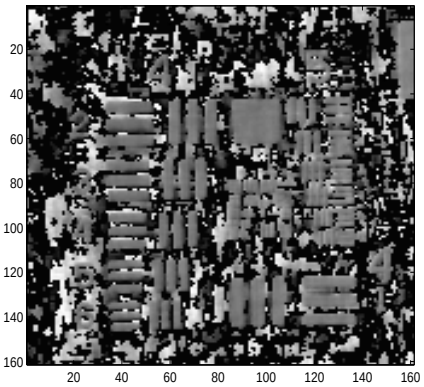
Figure 4.7 (a,b,c) show the phase reconstruction of groups 4 and 5 in the USAF chart with different exposure time. Fig. 4.7 (a) is reconstructed with low exposure time (400 μs), there are many speckle like noise (which is not the regular speckle noise in holography) at the reconstruction because of the shot noise. Fig. 4.7 (b) shows the reconstruction with longer exposure time (800 μs), although the noise speckles has now been reduced, it still exists. If we continue to increase the exposure time to 1500 μs, the reconstruction is shown in Fig. 4.7 (c), the noise speckles increases again, because of the saturation of the pixels. According to taking long exposure time we can reduce the effect of shot noise, however this is limited by the saturation. To overcome this restriction we introduce a technique that we term High Dynamic Range Holography (HDRH) [67, 68]. We take every hologram with different exposure time, for example in our case 800 μs, 1200 μs, 1600 μs, 2000 μs. Only pixel intensities between a 170 and 240 at different exposure times are kept and later normalized by the exposure time, then they are combined together to form a new hologram. This new hologram has no saturation points, also the shot noise level is reduced. From these results we see that to recover accurate phase information it is necessary to represent the digital signal with a higher number of bits which can be achieved using the HDRH approach as shown in Fig. 4.7 (d). The random phase fluctuations reduce significantly as the effective quantization depth representation of the digital signal is increased.



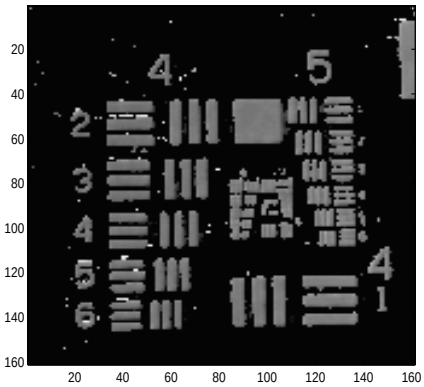
(a) Phase reconstruction with low exposure time (400 μs).



(b) Phase reconstruction with medium exposure time (800 μs).



(c) Phase reconstruction with long exposure time (1500 μs).



(d) Phase reconstruction using HDRH.

Fig. 4.7: Phase reconstruction with different exposure time and using HDRH technique.

4.5.3 Optimal power ratio between the reference and object arms

Here we analyze the optimal power balance between the object and reference waves based on the noise analysis above. Base on the argument above we examine Eq. (4.2) by considering the effect of detector shot noise. Hence N_n in Eq. (4.2), can be substituted into Eq. (4.6) and written in the following manner,

$$\Delta\psi_o(x,y) = \arg\left\{2H_K\cos[\Delta\psi_o(x,y)] + N_1 + N_3 + i\{2H_K\sin[\Delta\psi_o(x,y)] + N_2 + N_4\}\right\}. \tag{4.23}$$

From the equation above we see that, if there is no digital noise, $N_n = 0$, then the equation reduces to the correct answer $\Delta\psi_o(x,y)$. In order to minimize the effect of noise, we should keep the term H_K as large as possible relative to the noise term N_n.

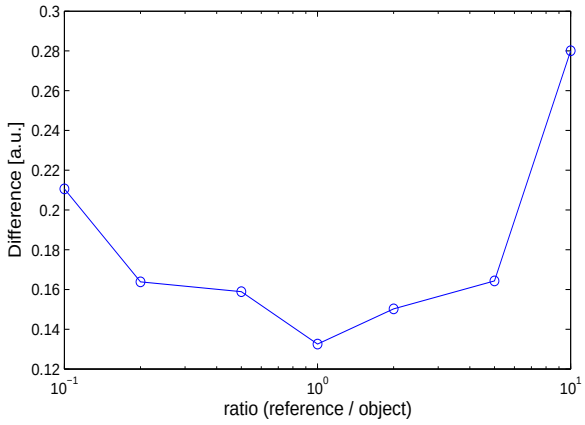


Fig. 4.8: Intensity difference between the measurement and the recovered one with different intensity ratio between reference wave and object wave

We define the ratio of the reference and the object wave is r_{ro},

$$r_{ro} = \frac{|A_r(x,y)|}{|A_o(x,y)|}. \tag{4.24}$$

To avoid the saturation, the following condition should be obeyed,

$$(|A_r(x,y)| + |A_o(x,y)|)^2 \leq I_{max} = 255. \tag{4.25}$$

For the purpose of reducing the shot noise, we take equal sign in the formula above. With the equation above, we have,

$$H_K = I_{max} \frac{2r_{ro}}{(1+r_{ro})^2}. \tag{4.26}$$

When $r_{ro} = 1$ we get a maximum value for the term H_K, which indicates that the intensity of reference wave and the object wave should be the same, to achieve the highest accuracy with minimum influence of shot noise, in keeping with the theoretical results presented here [53, 69, 70, 71]. In order to experimentally prove the argument above, we again contrast the recovered intensity and direct measured intensity as we did in Fig. 4.6 in Section 4.4.2, but with different power ratio. As shown in Fig. 4.8 the difference E at ratio $r_{ro} = 0.1$ is about 0.21, then r_{ro} is increased, the difference is reduced. When the power ratio goes to 1, the difference reaches the minimal, which agrees with our analysis. If we go on increasing the ratio, E goes larger again.

In this section we have analyzed the noise of the camera system and found its property meets the shot noise. We could either increase the laser power or the exposure time to reduce the effect of the detect noise. By using HDRH technique we increase the quantisation of the pixel value, which helps us to get a noiseless phase information. With the knowledge of the shot noise we show the optimal power

ratio between object and reference waves should be 1:1, then we have experimentally supported this argument.

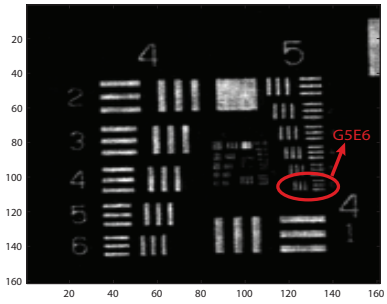
4.6 Experimental results

In this section, we will show the reconstruction results using the methods discussed above to reduce the effect of noise then try to get the resolution limit. We use a USAF chart as our object, by recognising the groups of three bars we could evaluate the quality and the image detail of the reconstruction. The CCD we used has 2056×2060 pixels, and the pixel size is $\gamma = 7.4$ μm, so the half sensor extend is $D \approx 1000 \times 7.4$ μm=7.4 mm. And the the coherent light source we used is HeNe Laser with wavelength $\lambda = 633$ nm.

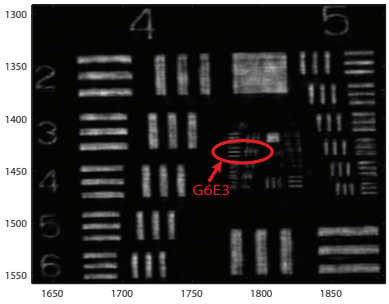
In the configuration above we take the distance between the sensor plane and the object plane with $z = 0.18$ m. As stated in Section 4.1, we have two factors that limit our resolution. One is related to the sensor extend with $D/\lambda z$ at 65 line/mm, another one is related to the pixel size with $1/\gamma$ at 135 line/mm, in our case the sensor extend limits our resolution. As shown in Fig. 4.9 (a), we can resolve the detail of the USAF chart with group 5, element 6 (G5E6 57 line/mm). We have approached the theoretical resolution limit at 65 line/mm.

We then reduce the distance between sensor and object to $z = 0.12$ m to get a higher resolution. At this condition we could get the resolution limit with 97 line/mm, corresponding G6E4 in USAF chart. Fig. 4.9 (b) shows the reconstruction when $z = 0.12$ m, we can resolve now 80.6 line/mm (G6E3), and it has approached to the resolution limit 97 line/mm. Note that the Nyquist frequency is the half of the sampling rate of a discrete signal processing system [72], in our case is only $1/2\gamma = 67.5$ mm^{-1}, corresponding G6E1, so the Nyquist limit does not limit the our holographic system at all.

From the experiment above we test and verify the resolution limit determined by the sensor extend by $D/\lambda z = 97$ line /mm. Now we are going to investigate another limitation concerning the pixel size. The resolution limit of the pixel size in the experiment above is $1/\gamma = 1/0.0074$ mm=135 line/mm, it is greater than the other one, so this limit will not affect our system. In order to make the pixel size resolution limit critical, we combine four pixel to one pixel, so the sensor extend resolution limit is still the same, however the other resolution limit becomes 135/2=67.5 line/mm, this resolution limit is now critical to our system. Fig. 4.9 (c) shows the reconstruction with pixel size 14.8 μm. We can resolve in this case about



(a) $z = 0.18$ m and pixel size $\gamma = 7.4$ μm



(b) $z = 0.12$ m and pixel size $\gamma = 7.4$ μm

(c) $z = 0.12$ m and pixel size $\gamma = 14.8$ μm

Fig. 4.9: Intensity reconstruction with different configuration z and γ

45.3 line/mm corresponding $G5E4$, the resolution has been decreased and is now limited by the pixel size.

4.7 A metrology application with digital holography

In the previous sections, we have talked about the digital holography and how the resolution of such systems are limited both theoretically and practically. In this section, metrology applications with digital holography are discussed.

Unlike traditional imaging systems, which capture only the intensity of the optical field, digital holography can measure both the intensity and phase of the optical field. The phase of optical field is related to the optical path length L_o, which is defined as the product of refraction index and the propagation distance. Consider a plane wave incident on a piece of glass as shown in Fig. 4.10. Because the refraction index of the glass is higher than air, the optical path length of the upper half is longer than the lower half, the optical path length difference is $\Delta L_o = (n_g - 1)c_g$, which c_g and n_g is the thickness and the refraction index of the glass. Therefore the wavefront of the upper half is delayed, which corresponds a phase delay of its optical field. With the help of digital holography, we can measure this phase

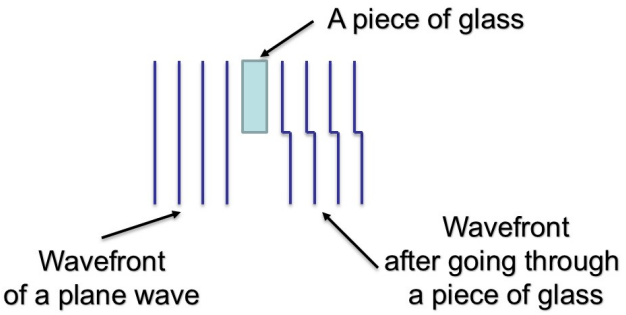


Fig. 4.10: Schematic depicting the changing of wavefront due optical elements.

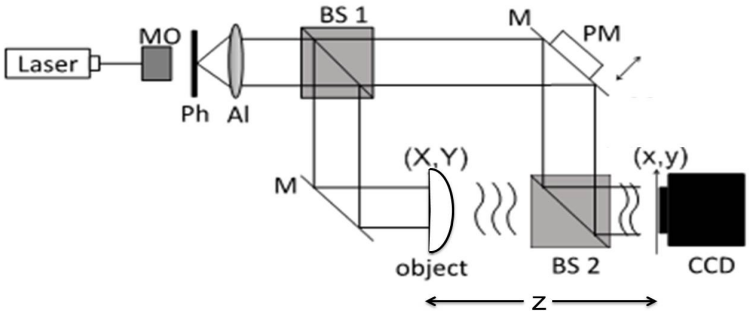


Fig. 4.11: Schematic depicting the changing of wavefront due optical elements.

difference for any given object. Note that the phase changing is very sensitive, any perturbation at the order of wavelength will cause changes in the hologram and thus can be measured. Therefore digital holography is a powerful technique for a range of metrology applications, such as vibration analysis [73, 74], deformation detection [75, 76], quantitative phase-contrast microscopy [3]. In the following section, we examine measuring a freeform surface using digital holography.

4.7.1 Freeform surface measurement

The holographical setup here is similar to the setup in Fig. 4.1. Instead of USAF chart used earlier, we have here the phase object, for example freeform lenses or other different kinds of lenses (Fig. 4.11). By performing the PSI in the manner described earlier, we reduce different kinds of error sources, and can reconstruct the phase changing due our phase object $\Delta\psi(X, Y)$. Assuming the laser we use has the wavelength λ in vacuum, the thickness and the refraction index at this wavelength of the phase object is $c_g(X, Y), n_g$. According to the phase changing $\Delta\psi(X, Y)$ we

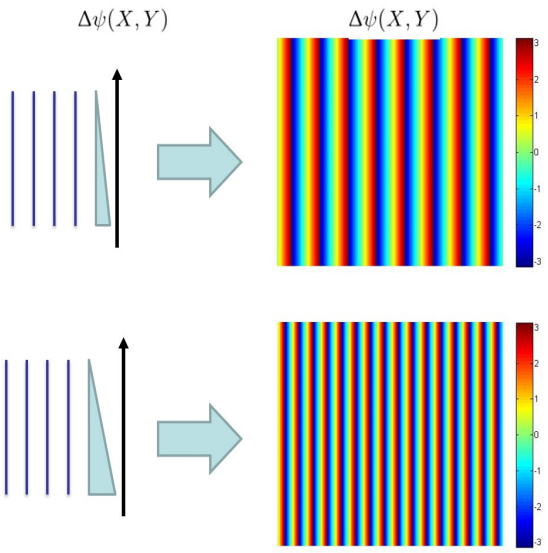


Fig. 4.12: Schematic depiction of the phase change when light goes through the glass prisms with different slopes.

could calculate the thickness of the phase object with,

$$c_g(X,Y) = \frac{\lambda}{2\pi(n_g-1)}\Delta\psi(X,Y). \tag{4.27}$$

As a holographic system, the resolution of the measurement is also limited, the resolution limit is reflected in how large the slope of the object surface can be measured. If the slope of the object surface is too large, which introduces high spatial-frequency components, such high spatial frequency components could not be recorded in the hologram due the finite extent of the sensor array or the convolution operation by the pixel size. In order to understand how the slope limits the measurement, we take two prisms with different slopes as example (Fig. 4.12). The phase changing due to the top prism with lower slope has a lower spatial frequency, which the prism with higher slope produces a higher spatial frequency, and this spatial frequency is limited by the resolution limit of the system discussed in Section 4.6.

Let's say the resolution limit of the system is R_{lim} line/mm, which corresponds the phase changing $\Delta\psi(X,Y) = 2\pi R_{lim}$ mm^{-1}. Substituting into Eq. (4.27), we have the measurable slope limit with,

$$\tan(\alpha) = \frac{\lambda R_{lim}}{n_g - 1} mm^{-1}. \tag{4.28}$$

If we have a system which is shown in Fig. 4.9 (b) with 97 line/mm resolution limit, and the light source is HeNe laser with $\lambda = 633$ nm, the refraction index of the object is $n_g = 1.5$, thus the maximal measurable slope of the object is about

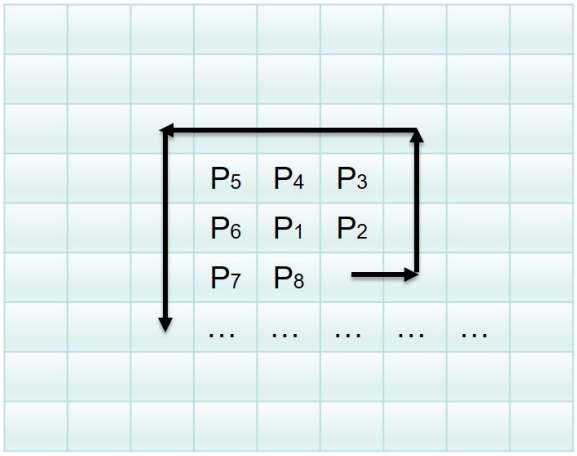


Fig. 4.13: Schematic depiction of the path-dependent phase unwrapping algorithm.

$\lambda R_{lim}/(n_g - 1) = 0.12$, and the maximal measured angle of the surface is about $\tan^{-1}(0.12) = 7°$. Although this holographical measurement has a high accuracy of the order of wavelength, the range of measurement is limited by the resolution limit of the system. To enlarge the range of measurement, we need decrease the pixel size of the sensor array and increase the sensor extent or decrease the distance between the object and sensor which was discussed in Section 4.6.

4.7.2 Phase unwrapping

4.7.2.1 Basic of a phase unwrapping process

Even if we can measure the phase change due an object with the demonstrated PSI technique, a problem remains: The phase function $\Delta\psi(X, Y)$ is wrapped over modulo 2π. This means the phase distribution is indefinite to an additive integer of 2π [77],

$$c_g(X, Y) = \frac{\lambda}{2\pi(n_g - 1)} \left[\Delta\psi(X, Y) + 2\pi n\right], \tag{4.29}$$

n is an integer. Therefore the thickness function contains 2π jumps at the position when $\Delta\psi(X, Y) = \pi$ or $\Delta\psi(X, Y) = -\pi$. In order to correct these 2π jumps to get a continuously varying phase, phase unwrapping is an essential step in the measurement.

Several unwrapping algorithms have been developed recent years [78, 79, 80]. In this section, we describe the so called path-dependent algorithm [77]. We considerate the two dimensional case. We imagine we have a wrapped phase distribution with a $N \times N$ pixels. To perform the path-dependent unwrapping we need find a path to pass through all the pixels. Usually we chose the center pixel as our starting point P_1, the second pixel P_2 is the pixel beside the first one P_1 as shown in Fig. 4.13. The next pixel P_3 goes surrounding the already unwrapped pixels and keep

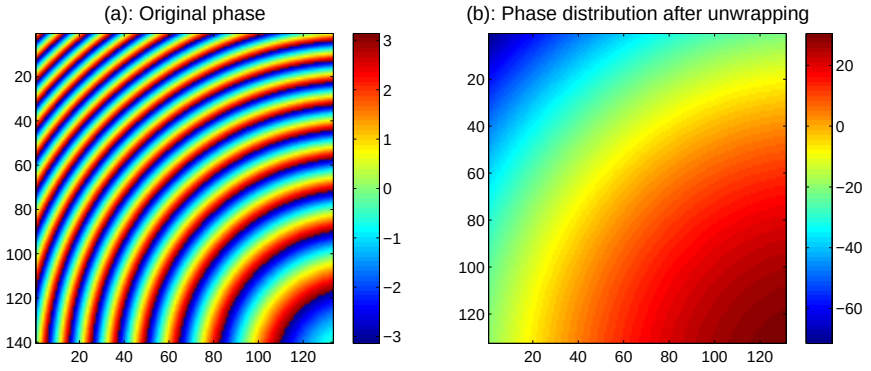


Fig. 4.14: (a): Original wrapped phase distribution; (b) Phase distribution after unwrapping.

going on until all the pixels are traversed. Similarly we reshape $\Delta\psi(X,Y)$ also as $\Delta\psi_1, \Delta\psi_2...$ corresponding the position of $P_1, P_2....$ At the begin of the traversing we make $P_1 = \Delta\psi_1$, the next unwrapped pixel P_{n+1} is calculated as following,

$$\begin{cases} P_{n+1} = \Delta\psi_{n+1} - \Delta\psi_n + P_n, & -\pi < \Delta\psi_{n+1} - \Delta\psi_n < \pi \\ P_{n+1} = \Delta\psi_{n+1} - \Delta\psi_n + P_n + 2\pi, & \Delta\psi_{n+1} - \Delta\psi_n < -\pi \\ P_{n+1} = \Delta\psi_{n+1} - \Delta\psi_n + P_n - 2\pi, & \Delta\psi_{n+1} - \Delta\psi_n > \pi \end{cases} \tag{4.30}$$

When all the N^2 pixels are all traversed, we will have the unwrapped phase distribution over the pixels P_1 to P_{N^2}. An unwrapping example is shown in Fig. 4.14. Fig. 4.14 (a) is the original wrapped phase distribution, and Fig. 4.14 (b) is the unwrapped phase distribution, which can be substituted into Eq. (4.27) to get the object surface profile.

4.7.2.2 Noise reduction before phase unwrapping

In the last section, we introduced the conventional path-dependent phase unwrapping algorithm, and found it works well for an ideal and noise-free simulation example. However in practice, the phase distribution suffers from different kinds of noise (see Sec. 4.3-4.5 for the noise analysis) (b), we can not get a "clean" phase distribution. In Fig. 4.15 (a) a wrapped phase distribution of a converging lens measured with digital holography is plotted, and Fig. 4.15 (b) is a zoomed in section of Fig. 4.15 (a). It can be seen, there are much noise at the boundary of the phase jumps. These noise effects are critical to the path-dependent phase unwrapping algorithm, and will make the unwrapping algorithm fail.

In order to reduce such noise we need to perform some type of smoothing or filtering operations. We propose here a new technique using Zernike fitting. We know that different orders of Zernike polynomials correspond different aberrations of the

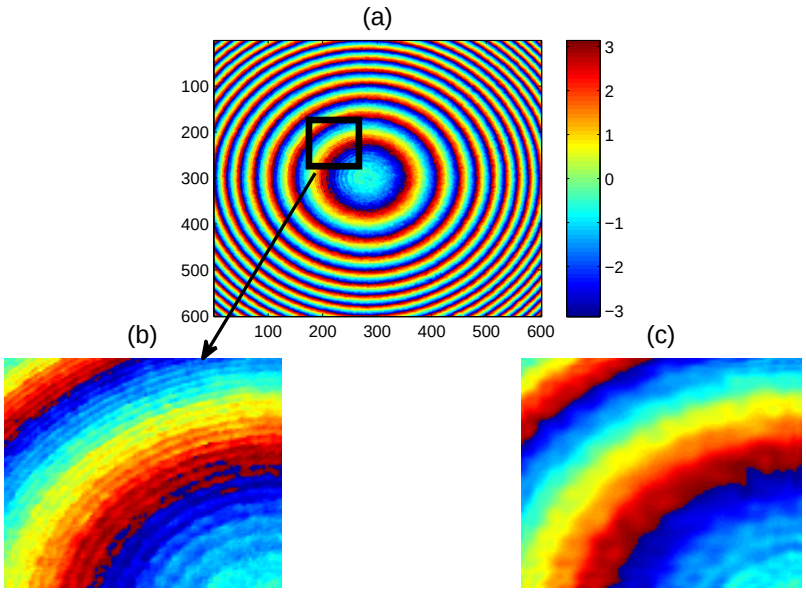


Fig. 4.15: (a): Wrapped phase distribution of a converging lens; (b) Partial enlarged drawing of (a); (c) Phase distribution after Zernike fitting.

lens, and the surface of lens can be decomposed into a finite number of Zernike polynomials. Therefore we use these Zernike polynomials to fit the reconstructed phase distribution by digital holography,

$$\exp(i\Delta\psi) = \sum_{n,m} \beta_n^m Z_n^m(\rho, \phi), \tag{4.31}$$

and use the acquired complex-valued Zernike coefficients β_n^m to regenerate the phase distribution. In this case, the noise is filtered by the Zernike fitting, we can then perform the unwrapping algorithm to get the thickness of object. Fig. 4.15 (c) shows the phase distribution of Fig. 4.15 (b) after Zernike fitting, the noise is significantly reduced, and the phase distribution is more smooth than the phase distribution before the filter.

It should be noted that the Zernike fitting process is based on the least mean squares method, its computational time depends on the pixel number of the fitted phase distribution. Therefore if we want to filter a phase distribution with 1000×1000 pixels, it will take over 10 mins to do the fitting process, and care must be taken to avoid running out of RAM when using Matlab. To reduce the computation time, we introduce a convenient processing step to improve the noise filtering operation. Instead of fitting the whole field $\exp(i\Delta\psi)$, we cut the field into small parts, and do the fitting for each individual section. We then combine all of these sections together. In Fig. 4.16 (a), the reconstructed phase distribution of Fig. 4.15 (a) is plotted. The

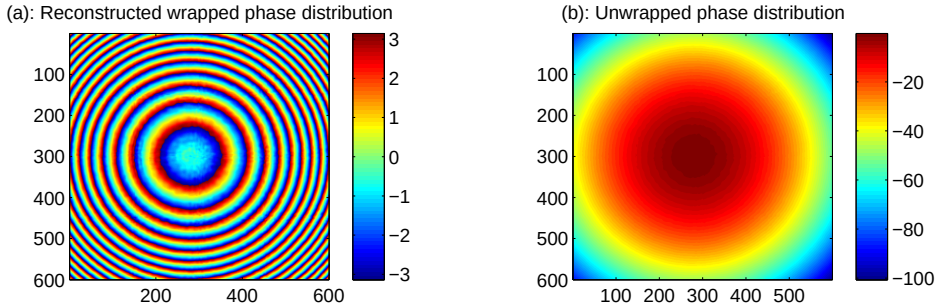


Fig. 4.16: (a): Reconstructed phase distribution of Fig. 4.15 (a) after Zernike fit. The original phase field is divided into 4×4 parts, and are combined together after the Zernike fit. (b): the unwrapped phase distribution of (a).

original phase field is divided into 4×4 sections and each section is fitted by Zernike polynomial with a maximal order of $n_{max} = 40$. After the fitting, the path-dependent unwrapping algorithm is performed, and the result is shown in Fig. 4.15 (b).

It must be emphasized that, this Zernike fitting technique is valid for the objects with smooth surface, which can be decomposed into finite orders of Zernike polynomials. Otherwise, this technique is not suitable, since a high number of Zernike terms are required for fitting the phase distribution, and may even introduce more noise into measurement result.

4.8 Conclusion

In this chapter we analyzed different kinds of practical experimental factors which reduce the accuracy of a Fresnel holographic setup with a PSI algorithm. These are: initial phase difference, PSI steps error, detection noise and vibration. In Section 4.3 we developed a mathematical model to account for vibrations and we analyzed the piezo motor jitter. It was shown that the piezo motor jitter will not affect the PSI system. For the initial phase difference, we could compensate them as shown in Section 4.4.1. In Section 4.4.2, we calibrated the PSI motor step size to correct the stepping error of the motor. In Section 4.5.1 for the detection noise we found in our investigation that the main noise of the CCD/CMOS array is the shot noise, whose SNR is proportional to the square of the intensity value of the pixel. By increasing the exposure time or using the high power laser we could reduce the speckle like noise in the reconstruction, however it is limited by the saturation of pixel. We introduced the HDRH technique to overcome this limitation which is shown in Section 4.5.2, then we have analytically and experimentally found that the best power ratio of reference and object wave should be 1:1.

With these methods above, in Section 4.6, we have corrected all these factor, and validated the two kinds of resolution limits of our holographic system: one is the diffraction limit with $D/\lambda z$, which is limited by the CCD aperture; the other one is related to the pixel size with $1/\gamma$, the detection sensor has filtered the signal with the pixel size, so the signal frequency higher than pixel frequency will be filtered. In order to achieve the high resolution imaging we need to considerate both these limitations. On one hand we could reduce the distance between the CCD array and the object, or increase the detection extent. The pixel size also needs to be taken into account.

In the last section, we discussed the metrology application with digital holography, especially a freeform surface measurement. Despite the high accuracy of the system, the range of measurement is still limited by the resolution limit. With the digital holographical system the phase modulation of the object can be measured, which can be related to the optical thickness of the object. However the measured phase is wrapped with 2π, hence an unwrapping algorithm is needed. The traditional path-dependent algorithm is a convenient way of unwrapping the phase, however it is not robust against noise. With the help of Zernike fit procedure, we can reduce the noise significantly and recover the surface profile of the object.

5 Iterative phase retrieval

In the last chapter, we discussed how the holographic technique can be used to measure the phase of optical field. In this chapter, another technique to estimate/measure the optical phase information, named iterative phase retrieval, is introduced and discussed. It differs from the holographic approach, which measures the phase using a reference wave. Iterative phase retrieval recovers the phase information from several intensity measurements made in axially displaced planes. In the following section, different phase retrieval techniques are introduced and compared. Based on the phase retrieval algorithm, we propose another technique to measure the phase.

5.1 Gerchberg-Saxton algorithm between the Fourier domains

In 1972, R. W. Gerchberg and W. O. Saxton proposed an iterative algorithm to recover the phase distribution of the optical field from two intensity measurements [81, 82].

Assuming there is an optical field $u(x) = |u(x)| \exp(i\psi_x)$ located in the front focal plane of a lens, see Fig. 5.1. The optical field in the back focal plane of the lens $\bar{u}(X) = |\bar{u}(X)| \exp(i\psi_X)$ corresponds to the optical Fourier transform of $u(x)$, see Chap. 5, P. 105 in Ref. [5],

$$\bar{u}(X) = \mathscr{F}\{u(x)\} \left\{ \frac{X}{\lambda f} \right\}. \tag{5.1}$$

With the help of optical digital sensor, we can measure the intensities in the front and back focal planes giving $|\bar{u}(X)|^2$ and $|u(x)|^2$. In order to get the phase information ψ_x and ψ_X in both planes, the Gerchberg-Saxton algorithm can be used.

The algorithm consists of the following six steps:
(1) Estimate the phase ψ_x^e of the $u(x)$, and do the Fourier transform of the signal;
(2) Replace the amplitude of the resulting Fourier transform with measured amplitude in the back focal plane $|\bar{u}(X)|$;
(3) Calculate the inverse Fourier transform of the signal after step 2;
(4) Replace the amplitude of the resulting inverse Fourier transform with measured amplitude in the front focal plane $|u(x)|$;

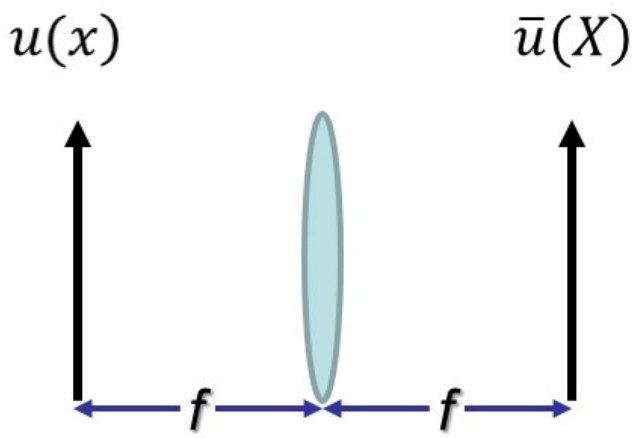


Fig. 5.1: Schematic depiction of the Fourier relationship in a lens system.

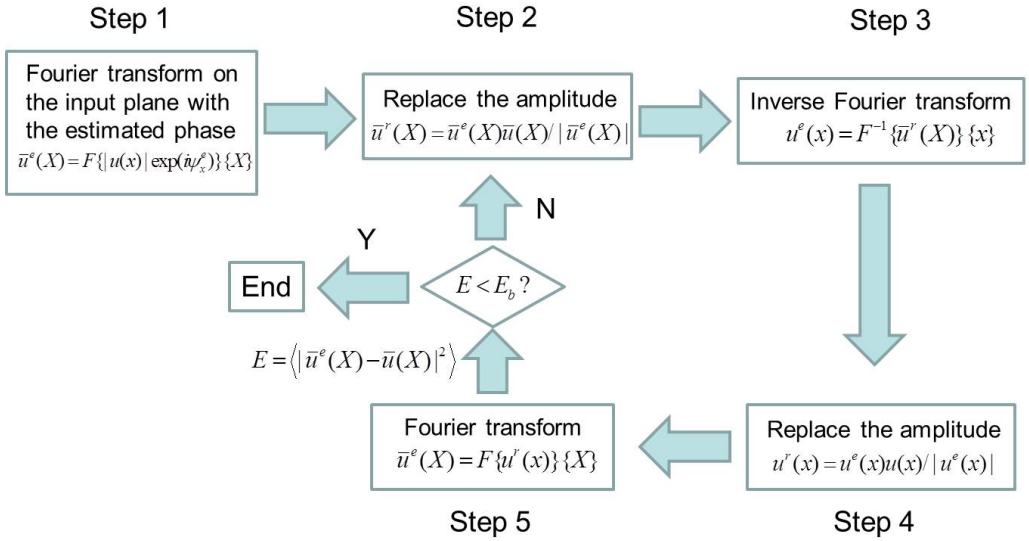


Fig. 5.2: Block diagram of the Gerchberg-Saxton algorithm.

(5) Calculate the Fourier transform of the signal after step 4;
(6) Redo step (2)-(4), until the difference between the estimated amplitude and measured amplitude decreases to a satisfactory level. The correspondent steps are depicted in Fig. 5.2. With this iterations algorithm, the phase of the optical fields are calculated.

In addition to measuring the phase information of the optical field, the Gerchberg-Saxton algorithm is also used in the design of diffractive optical elements [83, 84]. A diffractive optical element (DOE) is an optical element, which is used to shape or modulate an optical field to produce an output with a designed intensity profile. It is made of light transparent material, and works as a phase object. When the light propagates through the DOE, its phase is changed due the thickness of the DOE, see

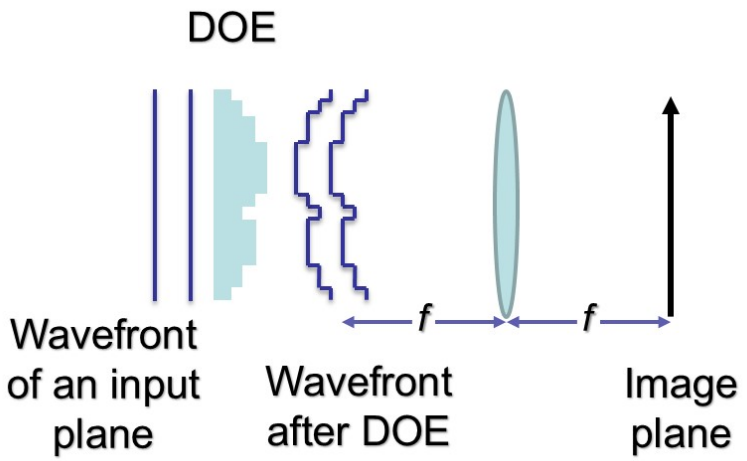


Fig. 5.3: Schematic depiction of a phase object (DOE).

Fig. 5.3. By appropriately fabricating the thickness of the DOE at different spatial locations, we can produce the designed optical field. Therefore we can generate arbitrary output intensity distributions in the back focal plane [85]. Assuming we have the intensity distribution of a plane wave illumination with $|u(x)|^2$, and we want to generate an intensity distribution with $|\bar{u}(X)|^2$. Substituting these two intensity values into the Gerchberg-Saxton algorithm, we can calculate the phase distribution ψ_x of the field $u(x)$. Then using Eq. (4.27), we calculate the thickness of the DOE in order to shape the beam intensity from $|u(x)|^2$ to $|\bar{u}(X)|^2$ in the image plane.

Note that in the iteration algorithm, both the forward Fourier and inverse Fourier transform operations are performed several times. To improve the speed of calculation, it is essential to use the FFT algorithm. However as mentioned in Chap. 2, the sampling intervals between two domains are different, in this case, their scaling factor depends on λ and f. Assuming the sampling interval of $u(x)$ is Δx with N elements $u(x_1), u(x_2)...u(x_N)$. After the FFT, we get the optical field $\bar{u}(x)$ with also N elements, however the sampling interval is now $\Delta X = \lambda f / N\delta x$.

In practice, we usually use the same sensor to measure the two intensity distributions, which means the sampling interval of the measured intensities $u(x)$ and $\bar{u}(X)$ are the same. However in the Gerchberg-Saxton algorithm, the sampling interval of the two intensity planes are different for the most case (except $\Delta x = \Delta X = \lambda f / N\delta x$). Therefore the measured intensities can not be directly substituted into the algorithm, which makes the algorithm not convenient in practice to retrieval the phase between two intensity measurement. In the next section, we will introduce another iterative algorithm, which is suitable for the intensity planes with the same sampling interval.

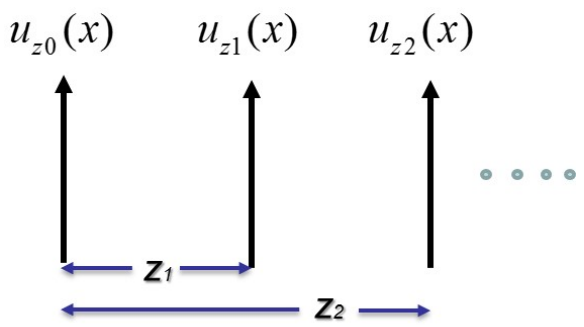


Fig. 5.4: Schematic depiction of the phase retrieval between Fresnel domains.

5.2 Phase retrieval using different Fresnel planes

In the last section, the intensity of the optical field is measured in an image plane and a Fourier plane. In this section, the optical field is measured in a series of free-space planes. The Fresnel transform is used to relate the fields in the different planes, see Fig. 5.4. For example the intensity distributions are captured in two different locations $|u_{z0}(x)|^2, |u_{z1}(x)|^2$. In this case, we can use the Fresnel transform to describe the free-space propagation of the optical field,

$$u_{z1} = \mathrm{FST}\{u_{z0}\}. \tag{5.2}$$

Similar to the Gerchberg-Saxton algorithm we also an iterative process as depicted in Fig. 5.2, however the Fresnel transform and the inverse Fresnel transform are used instead of the Fourier transform and inverse Fourier transform in Step 1,3 and 5 (by the inverse Fresnel transform we set $-z$ as the propagation distance instead of z in the Fresnel transform). To calculate this iterative process efficiently, we can use the FFT based propagation algorithms as outlined in Sec. 2.4. It is more convenient for us to calculate the Fresnel transform using the spectral method. The sampling intervals in both planes are identical, which means we can directly use the measured intensity signal in different planes, and substitute them into the iterative algorithm.

It must be emphasized that, as an iterative algorithm, the speed of convergence is important. In practice, the error decreases rapidly for the first few iterations, then more slowly for later iterations [81, 86]. In order to accelerate the speed of convergence, different kinds of techniques are used. For example instead calculating the optical field between two intensity planes, people found it converges more quickly when multiple intensity planes $u_{z0}, u_{z1}, u_{z2}...$ are used in the algorithm as shown in Fig. 5.4 [87]. Or we could modify the Gerchberg-Saxton algorithm and use the so-called hybrid input output algorithm to accelerate the convergence [88]. Another important factor that affects the speed of convergence is the initial estimated phase. If

the estimated phase approaches the real phase of the optical field, the algorithm will converge to the real phase in few iterations [89]. Otherwise, the algorithm converges not only very slow, but often converges to the incorrect local minimum, instead of the correct global minimum. In the next section, we propose a new technique to reconstruct the phase, in which we use the holographic system to make a good phase estimation and substitute the estimated phase into the phase retrieve technique.

5.3 Iterative reconstruction of digital holograms from three intensity measurements

Generally two methods are employed to capture and reconstruct in a digital hologram system:
(1): The Off-axis Fourier Spatial Filtering (OFSF) technique[44, 77], in which both the DC term, (i.e. average or zeroth order term) and twin image term are suppressed in the frequency domain;
(2): the Phase Shifting Interferometry (PSI) [7, 45] as we discussed in Chap. 4, in which a series of interference pattern images are acquired following controlled phase shifts of the reference beam.

The main advantage of the OFSF is that only one intensity pattern measurement (one image frame grab) is required. After a Fourier transform on the hologram, the real image term, the twin image term and the DC term are separated spatially in the frequency domain. After filtering the twin image term and the DC term, and use the inverse Fourier transform on the remaining real image term, we can reconstruct the object field. However, it has several disadvantages: (a) It is an off-axis technique rather than inline technique. Therefore it cannot be employed to capture and reconstruct in-line holograms; (b) It is difficult to capture holograms of speckle fields [77], so it cannot be used with strongly diffuse objects; and (c) Its resolution is limited by the reduced Nyquist rate achievable due to the use of an off-axis reference beam [4].

Compared with the OFSF, PSI is a more accurate digital holographic technique, and can be used to capture both in-line and off-axis digital holograms. However when applying PSI, the reference phase must be shifted in precise steps of a quarter of a wavelength between each of the (at least three) images that must be captured. Therefore: (1) The phase step accuracy requirement is a limitation of this technique [90, 91]; (2) The experimental setup for PSI is more complicated than in the OFSF case. In relation to (1) we note that there are several different ways to produce the required phase shifts. One is mechanically, using piezoelectric motors (e.g. in Chap. 4 in this thesis), another is by inserting appropriate wave

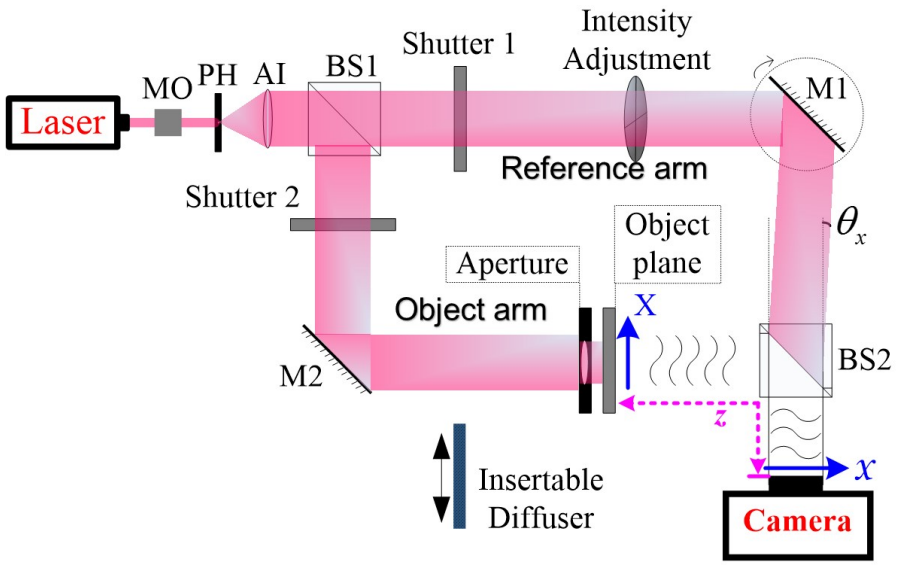


Fig. 5.5: Schematic depicting the measurement setup. MO, microscope objective; PH, pinhole; Al, achromatic lens; BS, beam-splitter; Mirror: M1 can be rotated to create both an in-line reference, (when $\theta_x = 0$, where θ_x is the angle of the reference wave), or an off-axis reference (when $\theta_x \neq 0$), and M2 is fixed; Note insertion of the diffuser generates a diffuse object field. The figure is taken from the Ref. [92].

plates (quarter, half etc.) into the reference beam. However precise piezoelectric motors can be expensive, while the typical accuracy of wave plates is not sufficiently high.

In this section, we propose a new digital holographic technique employing iterative computational algorithms [92], that overcomes the limitations of flexibility associated with the OFSF, and also overcomes the issues of for example piezoelectric motor cost associated with PSI. Furthermore the proposed system allows digital hologram data reconstruction in the case of diffuse objects, i.e. speckle fields. Finally, using the new method, which employs iterative phase retrieval algorithms [82, 93], twin image elimination is demonstrated.

5.3.1 Description of proposed technique

The setup of the proposed technique is shown in Fig. 5.5. Similar to the PSI setup, shown in Fig. 4.1. The laser light is collimated and separated into reference wave path and object wave path, then interfere again and captured by the CCD sensor. There are several differences: (i) Two shutters, one in the reference and object wave paths to block the reference wave or object wave at will; (ii) An aperture, placed before the object, to generate a clear boundary (sharp support) for the object image; and (iii) A diffuser, which can be inserted into the illumination beam before the aperture, (in order to randomize the illumination phase), to produce a diffuse

object field. Additionally, compared with a typical PSI setup, the proposed optical system is simpler because it lacks phase shift unit like wave plates or piezo motors.

As shown in Fig. 5.5 the space variable, x, refers to the coordinate in the camera plane, while X refers to the coordinate in the object plane. We assume that the original object wave-field in the object plane is $U_o(X,Y)$, while the object and reference wavefronts in the camera plane are $A_o(x,y)$ and $A_r(x,y)$. The resulting intensity distribution of the hologram is,

$$\begin{aligned} H(x,y) &= |A_o(x,y) + A_r(x,y)|^2 \\ &= |A_o(x,y)|^2 + |A_r(x,y)|^2 + A_o(x,y)A_r^*(x,y) + A_o^*(x,y)A_r(x,y) \\ &= |A_o(x,y)|^2 + |A_r(x,y)|^2 + A_{res}(x,y), \end{aligned} \tag{5.3}$$

$$A_{res}(x,y) = A_o(x,y)A_r^*(x,y) + A_o^*(x,y)A_r(x,y). \tag{5.4}$$

Note the terms $|A_o(x,y)|^2$ and $|A_r(x,y)|^2$ appearing in Eq. (5.3) can be measured by blocking either the reference or object wavefronts using the shutters, see Fig. 5.5. Once these two terms are removed, only the residual term, $A_{res}(x)$, remains, i.e. the sum of the real and twin image terms. If $A_{res}(x)$ is numerically back propagated to the object plane by applying an inverse Fresnel transform, then the resulting reconstructed image in the object plane will only contain contributions from the real and twin image terms. It should be noted that, in the object plane, the resulting real image contribution will be in-focus, while the twin image contribution is out of focus. Since the FST is a linear transform, the contributions from these two different terms (which are added together) can be mathematically considered separately,

$$\mathrm{FST}_{-z}\{A_{res}(x)\} = \mathrm{FST}_{-z}\{A_o(x,y)A_r^*(x,y)\} + \mathrm{FST}_{-z}\{A_o^*(x,y)A_r(x,y)\}. \tag{5.5}$$

In order to separate the real image from the twin image, we propose a practical numerical technique to eliminate the twin image contribution when:
(a) A well-defined image boundary (sharp support constraint) is present; and
(b) Applying iterative phase retrieval techniques.
It should be noted in relation to (a) that it is much easier to separate the real and twin image contributions if the original object image has a clear boundary [94]. In relation to (b), we note that, using standard/conventional iterative phase retrieval techniques it is well known that the object image phase can be reconstructed from two (or more) intensity measurements [93].

We note that an initial phase guess needs to be made when applying the conventional iterative phase retrieval techniques. This initial phase guess determines the

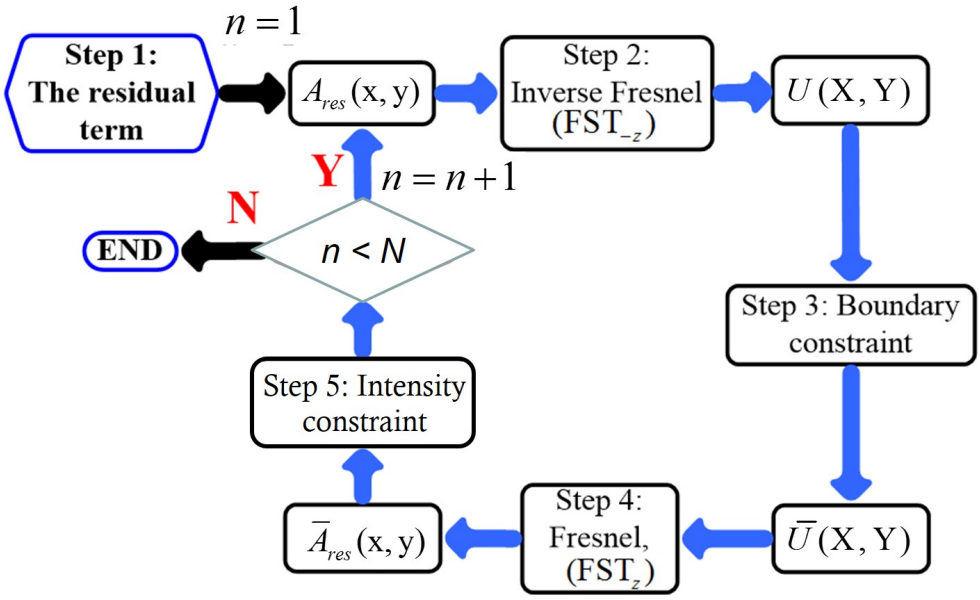


Fig. 5.6: Iterative phase retrieval algorithm flow chart, taken from the Ref. [92].

performance of corresponding phase retrieval algorithms:
(1) A "good" initial phase guess, (i.e. when the initial phase guess is close to the actual phase of the object), can improve the performance [89]; and
(2) A "bad" initial phase guess, (i.e. when the initial phase guess is different form the actual phase of the object), can significantly reduce the performance.

This poses one of the most serious limitations of iterative phase retrieval algorithms. However in the proposed method there is no need to make such initial phase guess. The reason for this is that in the proposed method we have a priori information, i.e. the residual term $u_{res}(x)$ (obtained from the experimental data, see Eq. (5.3)), for the starting phase distribution.

The flow chart for the iterative phase retrieval algorithm of the proposed method is shown in Fig. 5.6. In Fig. 5.6, N is the total (maximum) number of iterations, n indicates the nth iteration, $n = 1, 2...N$. In performing the numerical calculations a discrete algorithm, the spectral method is employed to approximate the continuous FST. The operation of the algorithm is as follows. To start we set, $n = 1$, and then: Step 1: The initial value of the residual term, $A_{res}(x, y)$, is calculated using the measured hologram, object, and reference intensities, $A_{res}(x, y) = H(x, y) - |A_o(x, y)|^2 - |A_r(x, y)|^2$.

Step 2: $A_{res}(x, y)$ is back propagated to the object plane employing the inverse FST, $U(X, Y) = FST_{-z}\{A_{res}(x, y)\}$.

Step 3: The field $U(X,Y)$ is replace by $\bar{U}(X,Y)$ which is generated by applying the boundary constraint in the object plane. $\bar{U}(X,Y)$ inside the boundary is identical to the corresponding field $U(X,Y)$. However at each iterative step the field outside the boundary are set to 0.

Step 4: $\bar{U}(X,Y)$ is forward propagated to the camera plane, giving a new estimate of the residual term $\bar{A}_{res}(x,y) = FST_z\{\bar{U}(X,Y)\}$.

Step 5: This $\bar{A}_{res}(x,y)$ is then modified using the known intensity constraint in the camera plane, $A_{res}(x,y) = \bar{A}_{res}(x,y) \times |A_o(x,y)|/|\bar{A}_{res}(x,y)|$. When $n < N$, the iteration value increments becoming $n+1$. Then the new value of $A_{res}(x,y)$ is used as the input for the next iteration, i.e. Step 2. When $n = N$, the algorithms stops.

5.3.2 Simulation and experiment results

The parameters used in our simulation are as follows: $\lambda = 633$ nm, $z = 162.55$ mm, the pixel size $\gamma = 7.4$ μm, and the number of pixels is 512×512. In the simulation, three cases are examined:

(a) A non-diffuse object having a small overlap area between the real and twin image contributions, (i.e. large reference beam angle $\theta_x \approx 0.363°$);

(b) A non-diffuse object having a large overlap area, (i.e. small reference beam angle $\theta_x \approx 0.073°$);

(c) A diffusely illuminated object, i.e. random phase at the camera by setting the phase of the object field randomly between 0 and 2π.

To perform the simulations we choose an original complex valued object field $U_o(X,Y)$. For the results presented the amplitude of $U_o(X,Y)$ is defined using a grayscale Lena image. A circular aperture of diameter 1.263 mm is applied to this image. We take $N = 20$ as the maximal number of iterations. The simulation results of these three cases are shown in Fig. 5.7.

The simulation results of Case. (a) is shown in the first columns of Fig. 5.7. For a large reference beam angle θ_x, the real image has only a small overlap with the twin image, see (a1). After the boundary constraint, most twin image is filtered. With this "good" initial phase guess and after 20 iterations, we have a relative good reconstruction (a2) compare to the reconstruction by PSI in (a3). If we reconstruct

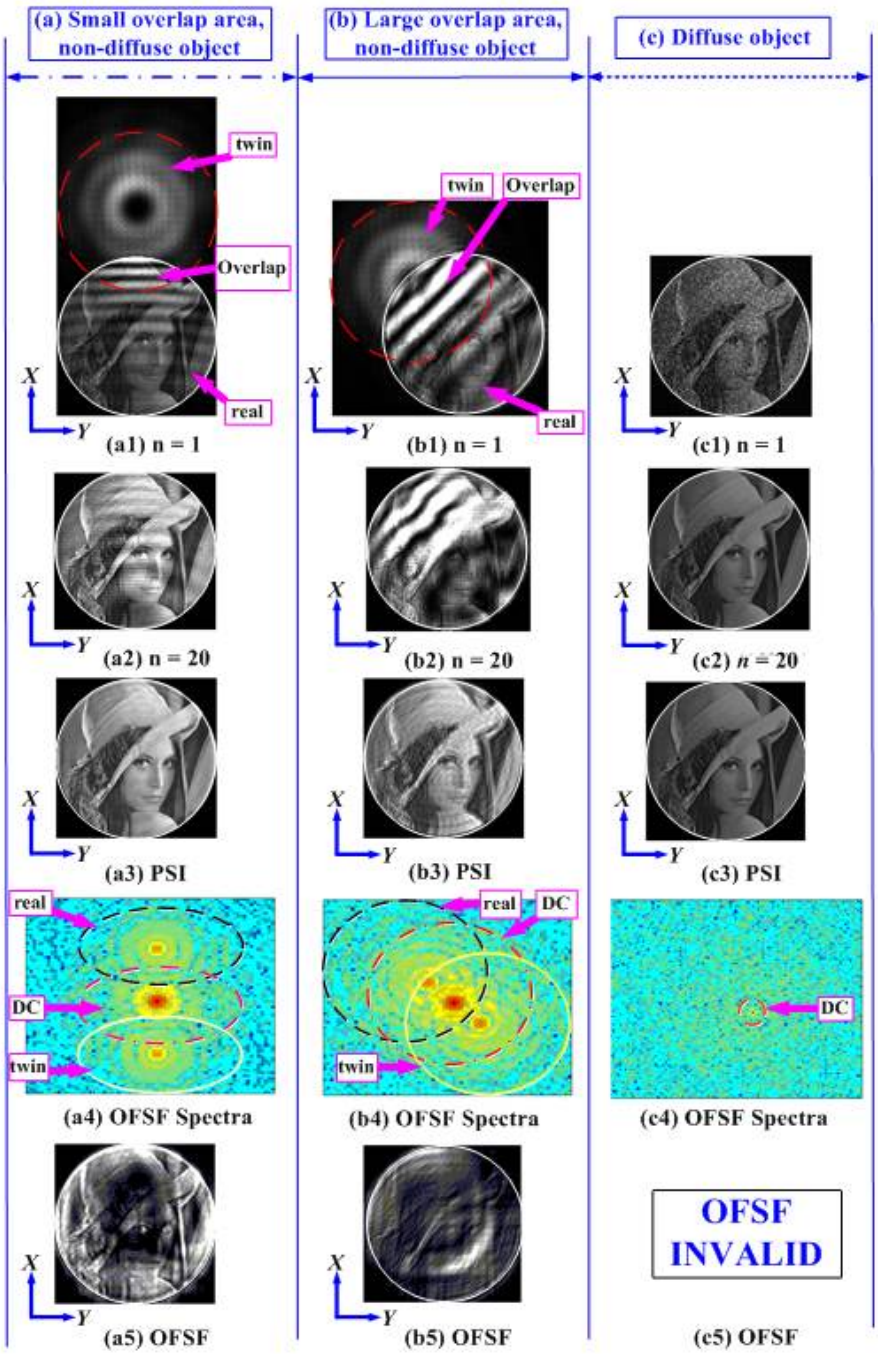


Fig. 5.7: Simulation results. The figures in the 1st, 2nd, and 3rd columns are reconstructed for: a non-diffuse object having: (a) Small overlap area, and (b) Larger overlap area; (c) diffuse object. The 1st and 2nd rows contain figures showing the reconstructed images obtained using the proposed method for $n = 1$ and $n = 20$ respectively. The figures in the 3th row are obtained using PSI. The figures in the 4th and 5th rows are obtained using the OFSF technique. The figure is taken from the Ref. [92]

the hologram using OFSF technique, the real image and twin image can not be well separated, see (a4), thus only a rough profile of the object can be reconstructed in (a5).

For Case (b) with a small reference beam angle, the simulation results are plotted in the second columns in Fig. 5.7. From (b1) we see that there is a big overlap area of the real image and twin image. After the boundary constraint, only a small part of twin image is eliminated, which means we have a "bad" initial phase guess. The reconstruction after 20 iterations is shown in (b2). Because of the "bad" initial phase guess, the reconstruction (b3) is failed. Also for the OFSF technique, see (b4) and (b5) the real image and twin image can not be separated in the frequency domain, and the reconstruction is also failed.

The last columns of Fig. 5.7 corresponds the Case (c), in which a diffuse object is simulated. In this case, the twin image is distributed over the whole object plane, only a small part of the twin image is overlapped with the real image. After 20 iterations, we found the reconstruction using our technique is as good as the reconstruction by using PSI, see (c2) and (c3). However for the OFSF technique, it is impossible to separate the real and twin images, the reconstruction is also poor.

From the simulation we found, the proposed technique works well if we have an off axis setup or a diffusely illumination, which means a "good" initial phase guess. Next, we will examine our technique experimentally. In the experiment we use a He-Ne laser with $\lambda = 633$ nm, an arbitrary chosen distance z=162.55 mm is applied (measured by a micrometer). The three intensities measurements were captured using a CCD camera (IMPERX IGV-B2020 with 2056×2060 pixels, having pixel size 7.4 μm).

As same as the simulation, the three cases discussed above were examined, Cases (a) and (b): Non-diffuse objects having (a) a large reference beam angle (off axis setup) or (b) a slightly reference beam angle (nearly in line setup); and Case (c): the diffuse object case. A standard USAF chart is used as the object with the diffuse object case being produced by inserting the diffuser into the beam illuminating the object as shown in Fig. 5.5.

The experiment results are demonstrated in Fig. 5.8. The performance of the proposed technique is similar to the simulation results. For Case (a) with a large reference beam angle, a small overlap between the real and twin images can be seen in (a1). After 20 iterations, the twin image is filtered, and we got a good reconstruction (a2) as well as the reconstruction by PSI (a3) and OFSF (a5). For Case (b) with a nearly inline setup, the twin image and the real image are overlapped, we got a bad reconstruction as same as the OFSF (b5). For the diffuse case (c), the reconstruction of our technique (c2) performs as well as the reconstruction by PSI (c3), however the result with OFSF fails due the overlap of real and twin images in the frequency

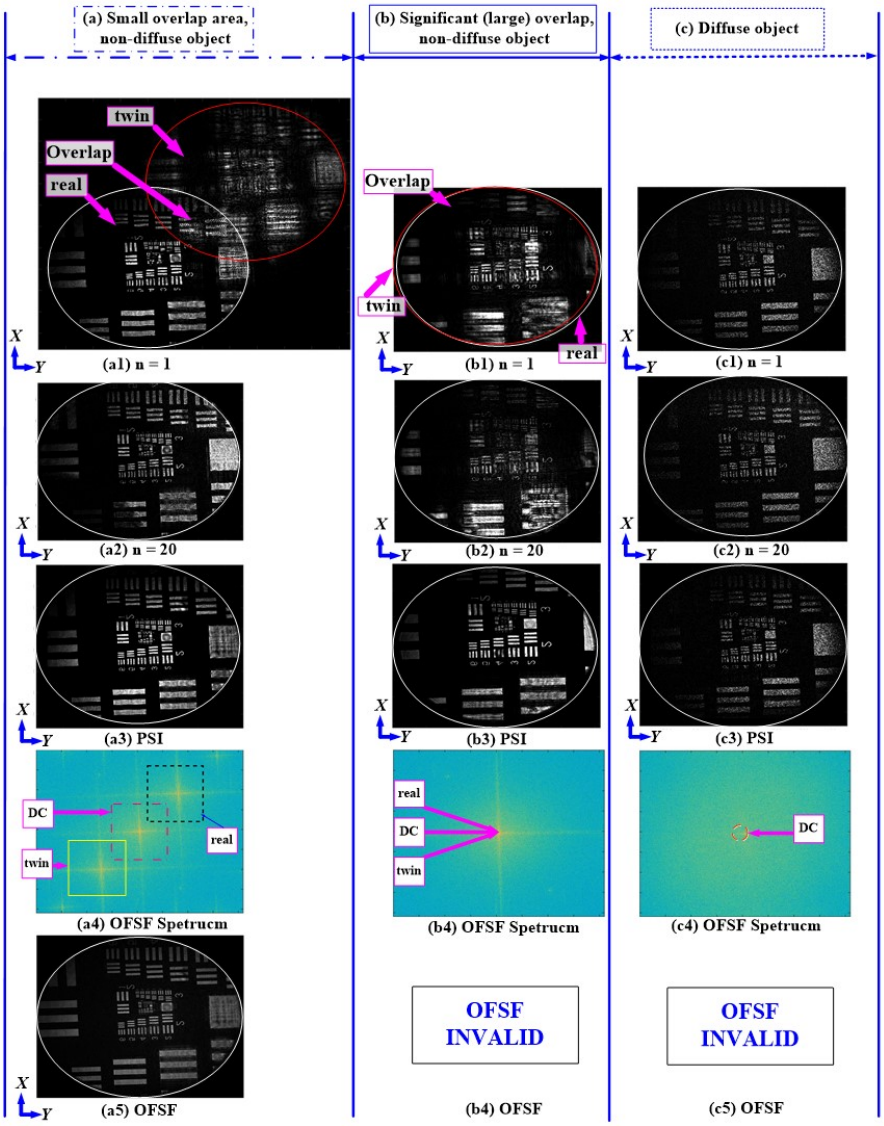


Fig. 5.8: Experimental results. Comparison of the reconstructed image intensities obtained by the proposed method, PSI and OFSF for three cases: A non-diffuse object having (a) a small overlap area (large off-axis reference angle of incidence) and (b) a significant (large) overlap area; (c) a diffuse object. The figure is taken from the Ref. [92]

domain. The experimental results coincide with the simulation results for the three different cases.

In comparison to previously used PSI, OFSF methods, there are several benefits of the proposed method. First, compared with the PSI, no expensive piezo motor or quarter wave plates are required, resulting in a much simpler experimental setup. Second, the proposed method works as well as the OFSF for the large reference angle case, and also works for the diffuse case, for which the OFSF fails.

5.4 Conclusion

In this chapter, we have introduced the iterative phase retrieval technique, which estimates the optical phase by measuring the intensity of the field in two or more planes. Compared to the digital holography, phase retrieval doesn't need reference wave and therefore has a relative simple setup. However the accuracy of the phase retrieval depends on the initial phase guess. If a "bad" initial phase is chosen, the phase retrieval algorithms converges slowly and usually converges to a local minimum rather than the global minimum solution. Hence the phase distribution is often incorrectly estimated. Therefore, we proposed a technique, combining a holographic system with a phase retrieval technique. The reconstructions with the proposed method are simulated and experimentally tested.

6 Diffraction under partially coherent illumination

6.1 Introduction

In the previous chapters, we have assumed that the light in our optical system was monochromatic and totally coherent. An example of a coherent beam is a well stabilised laser. Incoherent beam are more common in daily life and include sunlight and the light from LEDs. Between these two limiting cases coherent and incoherent, we have what we refer to as the partially coherent beam. The partially coherent beam has received a lot of attention in recent decades, like the long range communication, speckle reduction [95, 96]. Hence it is important to be able to simulate the propagation of partially coherent beam efficiently and accurately.

For a fully coherent beam, we can use diffraction integral, for example the Fresnel transform, to calculate its propagation. The phase of an optical field varies rapidly over time. For a coherent field, while the phase may vary rapidly, all parts of the field vary in unison. For a partially coherent field this is no longer true. And different points of the field (in a given plane) vary randomly and rapidly respect to each other. So for the partially coherent beam, it is not possible to use a static diffraction integral to describe the intensity that would be measured by a detection. To model this partially coherent case we can use the mutual coherence function to describe the partially coherent beam, and calculate the propagation of this mutual coherence function using a 4D mutual coherence integral, see Chap 5.4 in Ref. [40]. However, as for the Fresnel transform calculations, for the most cases, there are no analytical solutions for the mutual coherence integral and we must do the integral numerically. As a 4D integration, it is challenging to solve it using the methods that were implemented in Chap. 2.

In this chapter, we decompose the dynamic partially coherent field into a temporal sequence of static spatially coherent random fields. By propagating these spatially coherent random fields separately and adding them together on an intensity basis, we can model the output of the 4D mutual coherence integral. In order to verify the proposed technique, we simulate the Thompson-Wolf experiment and compare the simulated diffraction pattern with these theoretical results. At the end, we use

the proposed technique to simulate the performance of DOE under partially coherent illumination, some interesting results are found and discussed.

6.2 The Thompson-Wolf experiment

In 1957, B. J. Thompson and E. Wolf developed an experiment, in order to investigate the partial coherence of a luminous source [97]. In this section, we will review their experiment, introduce and characterize some fundamental concepts of the partially coherent beam.

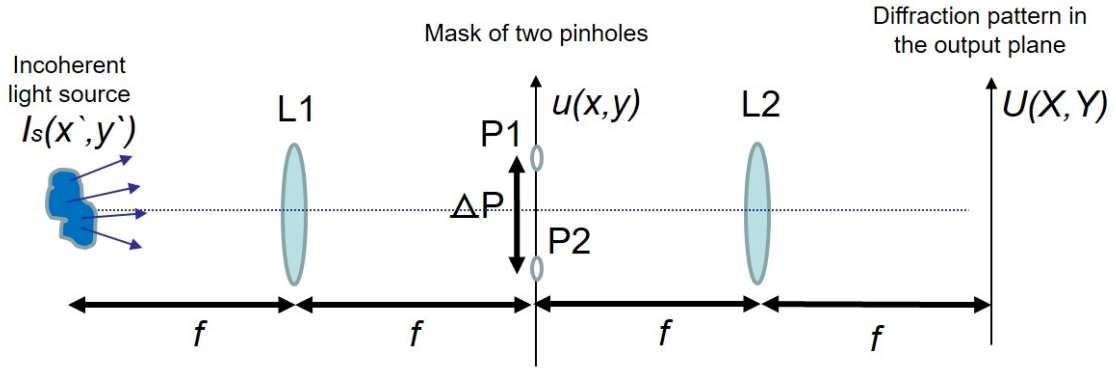


Fig. 6.1: Illustrating the setup of the Thompson-Wolf experiment

Consider a finite quasi-monochromatic source, locates in the front focal plane of a lens L1. In the back focal plane of L1 we have a diffracting mask of two pinholes P1 and P2. Its Fourier pattern is captured in the back focal plane of another lens L2. Let $u(x_1, y_1; t)$ and $u(x_2, y_2; t)$ be the complex optical field at points P1 and P2, at time t, and $\Gamma(P1, P2)$ is known as the mutual coherence function,

$$\Gamma(P1, P2) = \langle u(x_1, y_1; t) u^*(x_2, y_2; t) \rangle \,, \tag{6.1}$$

where the angular brackets "$\langle\rangle$" is the ensemble average operation over time. Know that the light source plane and the pinhole plane have a Fourier relationship, and we can write the mutual coherence function as the Fourier transform of the light source $u_s(x', y'; t)$,

$$\begin{aligned} &\langle u(x_1, y_1; t) u^*(x_2, y_2; t) \rangle = \\ &\left\langle \iint u_s(x'_1, y'_1; t) e^{-\frac{i2\pi(x'_1 x_1 + y'_1 y_1)}{\lambda f}} dx'_1 dy'_1 \iint u_s^*(x'_2, y'_2; t) e^{\frac{i2\pi(x'_2 x_2 + y'_2 y_2)}{\lambda f}} dx'_2 dy'_2 \right\rangle . \end{aligned} \tag{6.2}$$

In the integration above, only the complex field of light source $u_s(x', y'; t)$ is changing over time, so we rewrite the Eq. (6.2) as,

$$\langle u(x_1, y_1; t)u^*(x_2, y_2; t)\rangle = \iint\iint \left\langle u_s(x'_1, y'_1; t)u^*_s(x'_2, y'_2; t)\right\rangle e^{-\frac{i2\pi(x'_1x_1+y'_1y_1)}{\lambda f}} e^{\frac{i2\pi(x'_2x_2+y'_2y_2)}{\lambda f}} dx'_1dy'_1dx'_2dy'_2. \quad (6.3)$$

Note that an incoherent light source is characterised as Dirac delta correlated, thus we have,

$$\left\langle u_s(x'_1, y'_1; t)u^*_s(x'_2, y'_2; t)\right\rangle = u_s(x'_1, y'_1; t)u^*_s(x'_2, y'_2; t)\delta(x'_1 - x'_2)\delta(y'_1 - y'_2). \quad (6.4)$$

Substituting Eq. (6.4) into Eq. (6.3) and Eq. (6.1), we reduce the 4D integral into a 2D integral with,

$$\Gamma(P1, P2) = \iint I_s(x', y')e^{-\frac{i2\pi[x'(x_1-x_2)+y'(y_1-y_2)]}{\lambda f}} dx'dy', \quad (6.5)$$

which $I_s(x', y') = u_s(x'_1, y'_1; t)u^*_s(x'_2, y'_2; t)$ is the intensity of the light source. Note that the mutual coherence function between P1 and P2 depends only on their spatial distance $\Delta P = x_1 - x_2$ in x direction, and $\Delta Q = y_1 - y_2$ in y direction, therefore $\Gamma(P1, P2)$ can also be written as,

$$\Gamma(P1, P2) = \Gamma(\Delta P, \Delta Q) = \iint I_s(x', y')e^{-\frac{i2\pi(x'\Delta P+y'\Delta Q)}{\lambda f}} dx'dy'. \quad (6.6)$$

The mutual coherence function in the pinhole plane is the Fourier transform of the intensity pattern of the incoherent light source. In the Thompson-Wolf experiment, they had a incoherent light source with uniform intensity in a circular aperture of a radius r'. According to Eq. (6.6), the mutual coherence function in this case is the Fourier transform of a circular aperture, which is a Jinc function,

$$\Gamma(\Delta P, \Delta Q) \propto Jinc\left(\frac{2\pi r'\sqrt{\Delta P^2 + \Delta Q^2}}{\lambda f}\right) \quad (6.7)$$

$$Jinc(x) = \frac{J_1(x)}{x}. \quad (6.8)$$

Note that the mutual coherence function here is circularly symmetric, so we have,

$$\Gamma(\Delta P, \Delta Q) = \Gamma(r) \propto Jinc\left(\frac{2\pi r'r}{\lambda f}\right), \quad (6.9)$$

$$r = \sqrt{\Delta P^2 + \Delta Q^2}. \quad (6.10)$$

By normalizing the mutual coherence function, we have the complex degree of the coherence $\mu(P1, P2)$ with,

$$\mu(P1, P2) = \mu(r) = \frac{\Gamma(P1, P2)}{[\Gamma(P1, P1)\Gamma(P2, P2)]^{1/2}} = 2Jinc\left(\frac{2\pi r' r}{\lambda f}\right). \tag{6.11}$$

Note that this complex degree of the coherence μ characterizes the correlation between the two beams, and it amplitude has to be between 0 and 1,

$$0 \leq |\mu(P1, P2)| \leq 1. \tag{6.12}$$

Consider the light source has a very small radius $r' \to 0$, as a point source. Then even for a large ΔP between two pinholes, we have a complex degree of the coherence $\mu(P1, P2) \to 1$. In this case, we have a fully coherent illumination, the diffraction pattern of these two pinholes at their Fourier plane is then,

$$I_t(X, Y) = 2\cos\left[\frac{2\pi}{\lambda f}(\Delta PX + \Delta QY)\right], \tag{6.13}$$

we have a maximal visibility v_m of 1, which is defined as [97],

$$v_m = \frac{Imax - Imin}{Imax + Imin}. \tag{6.14}$$

If the light source is no longer a point source and has a large spatial extent r'. The complex degree of the coherence is then decreased between 0 and 1. For this instance, we have a partially coherent illumination, and the diffraction pattern is then,

$$I_t(X, Y) = (1 + |\mu(P1, P2)|)\cos\left[\frac{2\pi}{\lambda f}(\Delta PX + \Delta QY)\right], \tag{6.15}$$

the visibility v_m is decreased as well as the complex degree of the coherence $\mu(P1, P2)$, $v_m = \mu(P1, P2)$. In practice, the pinhole can not be infinite small and has a diameter of d_i. So we modify the Eq. (6.15) by considering the extent of pinhole (see Ref. [40], Eq. (5.2-44) for detail),

$$I_t(X, Y) = \left[\frac{J_1\left(\frac{\pi d_i}{\lambda f}\sqrt{X^2 + Y^2}\right)}{\frac{\pi d_i}{\lambda f}\sqrt{X^2 + Y^2}}\right]^2 \left\{1 + |\mu(\Delta P, \Delta Q)|\cos\left[\frac{2\pi}{\lambda f}(\Delta PX + \Delta QY)\right]\right\}. \tag{6.16}$$

At this point, we have derived the analytical express of the diffraction pattern in the Thompson-Wolf experiment. In the next section, we will simulate the Thompson-Wolf experiment and compare the results with this analytical express.

6.3 Wave optics simulation of partially coherent beams

In the last section, we reviewed the Thompson-Wolf experiment and introduced some important concepts of the coherence theorem. In this section, we will investigate how partially coherent beam propagates and how to calculate its propagation efficiently.

In our approach, instead of solve the 4D integration directly, we generate enough more random complex fields $u(x, y; t)$, with the same mutual coherence function Γ. Then we propagate these random spatially coherent fields, and have their complex field in the output plane $U(X, Y; t)$. By adding adding them together on an intensity basis, we have the intensity pattern I_t^N in the diffraction plane of the two pinholes.

In order to better understand and validate the proposed technique, we simulate the Thompson-Wolf experiment introduced in the last section and compare its intensity pattern I_t^N in the output plane with its theoretical pattern in Eq. (6.16).

Consider an incoherent light source, see Fig. 6.1, whose intensity is uniform in a circular aperture of radius r',

$$\begin{cases} I_s(x', y') = 1, x'^2 + y'^2 \leq r'^2) \\ I_s(x', y') = 0. x'^2 + y'^2 > r'^2) \end{cases} \tag{6.17}$$

And we can write the complex field of this incoherent light source as,

$$u_s(x', y'; t) = \sqrt{I_s(x', y')} \exp^{i\theta(x', y'; t)} . \tag{6.18}$$

Because the light source is incoherent, the phase of the light source $\theta(x', y'; t)$ at any instance t is uncorrelated with each other between 0 and 2π. This complex random field $u_s(x', y'; t)$ propagates through the lens L1, and generates a speckle field $u(x, y; t)$ in the back focal plane of L1, which is characterized as the Fourier transform of $u_s(x', y'; t)$,

$$u(x, y; t) = \mathscr{F}\{u_s(x', y'; t)\} \left\{ \frac{x}{\lambda f}, \frac{y}{\lambda f} \right\} . \tag{6.19}$$

According to Eq. (6.11), the normalised mutual coherence function of $u(x, y; t)$ is a Jinc function,

$$\mu(x_1, y_1; x_2, y_2) = \mu(\Delta x, \Delta y) = \mathscr{F}\{I_s(x', y')\} \left\{ \frac{\Delta x}{\lambda f}, \frac{\Delta y}{\lambda f} \right\} = 2Jinc \left[\frac{2\pi r'(\Delta x^2 + \Delta y^2)}{\lambda f} \right] , \tag{6.20}$$

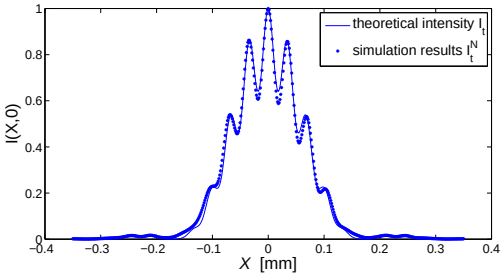


Fig. 6.2: Comparison of the theoretical intensity distribution $I_t(X,0)$ to the simulated results $I_t^N(X,0)$ with the degree of coherence $\mu = 0.2$.

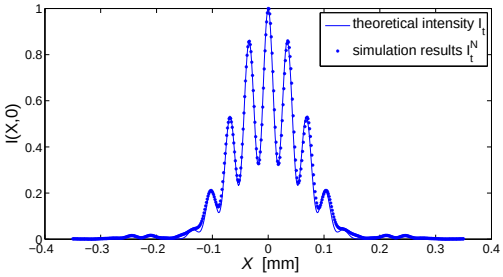


Fig. 6.3: Comparison of the theoretical intensity distribution $I_t(X,0)$ to the simulated results $I_t^N(X,0)$ with the degree of coherence $\mu = 0.5$.

where $\Delta x = x_1 - x_2$ and $\Delta y = y_1 - y_2$. And we can manipulate the complex degree of the coherence by choosing different r' according to Eq. (6.20). By multiplying the generated random field $u(x,y;t)$ with the pupil function $P_p(x,y)$ of the pinholes and performing the Fourier transform of it, we have the complex field $U(X,Y;t)$ at the output plane,

$$U(X,Y;t) = \mathscr{F}\{u(x,y;t)P_p(x,y)\}\left\{\frac{X}{\lambda f}, \frac{Y}{\lambda f}\right\}, \tag{6.21}$$

$$\begin{cases} P_p(x,y) = 1, (x,y) \text{ inside the pinhole area} \\ P_p(x,y) = 0. (x,y) \text{ outside the pinhole area} \end{cases} \tag{6.22}$$

And we have thus the intensity pattern $I_t^N(X,Y)$ at the output plane,

$$I_t^N(X,Y) = \left\langle |U(X,Y;t)|^2 \right\rangle. \tag{6.23}$$

In our simulation, we have $\lambda = 700$ nm, $d_i = 1$ cm, $\Delta P = 4$ cm, $\Delta Q = 0$ cm, $f = 300$ mm. By properly choosing the σ in Eq. (6.20), we take three different configurations with, $\mu(\Delta P, \Delta Q) = 0.2, 0.5, 0.8$ (We use Eq. (6.20) to chose correspondent radius r' of the incoherent light source). For every configuration 200 random speckle fields $u(x,y;t)$ are generated using Eq. 6.19, and substituted into Eq. (6.21) and Eq. (6.23). The intensity distribution $I_t^N(X,Y)$ along X axis at the output plane are calculated and plotted in Fig. 6.2-6.4. A high agreement can been seen between the simulation and the theoretical results.

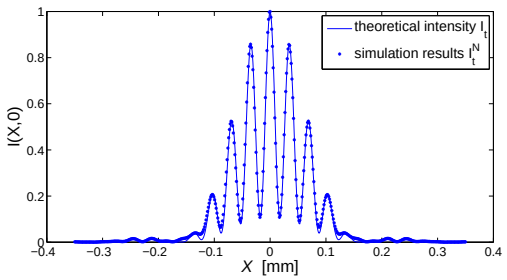


Fig. 6.4: Comparison of the theoretical intensity distribution $I_t(X,0)$ to the simulated results $I_t^N(X,0)$ with the degree of coherence $\mu = 0.8$.

6.4 Simulation of the DOE diffraction pattern with the partially coherent beam

As we introduced in Chap. 5.1, diffractive optical element (DOE) uses a thin micro structure pattern to alter the phase of the light that passes through it. This phase pattern, once properly designed, can manipulate the light so that almost any desired intensity profile can be produced. Examples include beam shaping and splitting or generating cartoon figures for projection system. To design a DOE with a particular designed output intensity profile, the Iterative Fourier Transformation Algorithm (IFTA) after Gerchberg and Saxton [81] (Detailed in Chap. 5) is often used. With this algorithm we can calculate a phase distribution that can be etched onto a DOE to produce a designed intensity profile in the output domain.

Note that the DOEs are designed with the assumption that the light is coherent. The coherence of the illumination light plays an important role of the performance of the DOEs. We expect that the diffraction pattern loses its clarity as the light source becomes increasingly incoherent. In this section, we are interesting in examining this effect in more detail using numerical simulations. We show how partially coherent beam affects the performance of the DOE, and provide design guidelines to ensure that the DOE operates within specific limits under partially coherent illumination.

We consider a DOE, which is designed to transform an incident plane wave into a intensity profile given by a series of rings [target intensity profile, see Fig. 6.5 (a)]. Using the IFTA method proposed by Gerchberg and Saxton [81], we designed our DOE phase pattern with 2000×2000 features, each feature has an extent of 0.01 mm. If we use a laser illumination on the designed DOE, we expect that we can have a profile with the sharp edges, which is shown in Fig. 6.5 (b). If the DOE is illuminated by a partially coherent beam instead, we expect that its diffraction pattern will get blurred. In this simulation, we use an incoherent light source with a

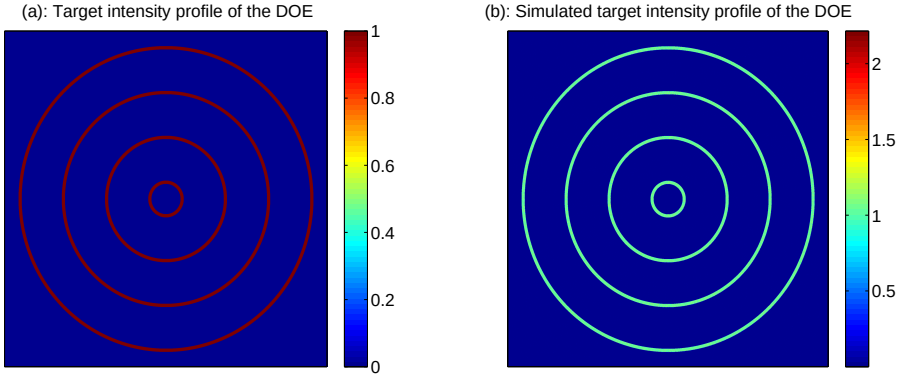


Fig. 6.5: Intensity profile of the DOE diffraction pattern: (a) target intensity profile, (b) simulated intensity profile under totally coherent illumination.

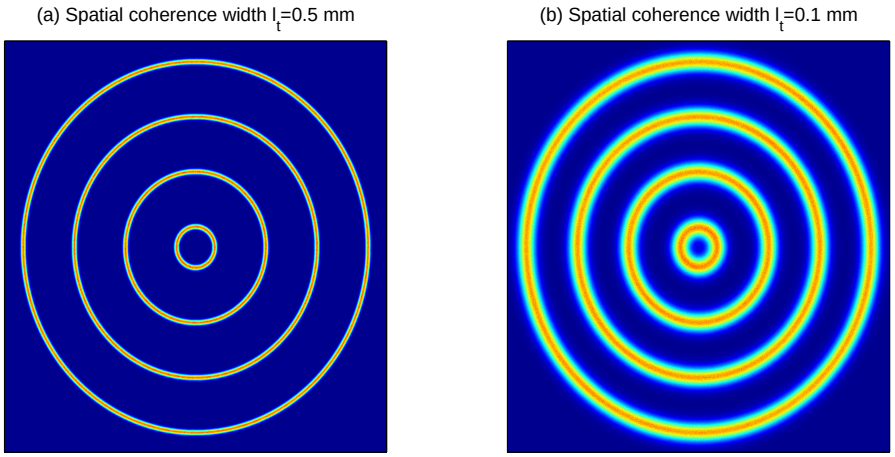


Fig. 6.6: Intensity profile of the DOE diffraction pattern under partially coherent illumination: (a) $l_t = 0.5$ mm, (b) $l_t = 0.1$ mm.

Gaussian intensity profile instead of the circular aperture,

$$Is(x', y') = \exp\left(-\frac{x'^2 + y'^2}{2\sigma^2}\right), \tag{6.24}$$

which σ is the standard deviation. In this case, the degree of coherence function μ is also Gaussian. Using the method we proposed above, we calculate the DOE diffraction patterns under spatially coherent illumination with different spatial coherence width l_t, which is defined as the half width or standard deviation of the Gaussian complex degree of the coherence μ,

$$\mu(l_t) = \exp\left[-\frac{2\pi^2\sigma^2 l_t^2}{\lambda^2 f^2}\right] \approx 0.606, \tag{6.25}$$

$$l_t = \mu^{-1}(0.606). \tag{6.26}$$

The simulation results are shown in Fig. 6.6. The intensity patterns are calculated under partially coherent illumination with the spatial coherence width 0.5 mm and 0.1 mm. The intensity pattern with the large spatial coherence width (Fig. 6.6 (a)) is blurred compared to the totally coherent illumination (Fig. 6.5 (b)), while the intensity pattern with the small coherence width (Fig. 6.6 (b)) is much more blurred. To quantify this blurred effect, we use the Michelson visibility v_m, which is defined in Eq. (6.14), as a criterion to evaluate the diffraction pattern of the DOE. In Fig. 6.6 (a), the intensity pattern is similar to Fig. 6.5 (b) due the large spatial coherence width, we have a high visibility of 0.9174. However for the Fig. 6.6 (b), which has the illumination with a small spatial coherence width, the visibility is reduced to 0.5659. In order to investigate the effect of partially coherent illumination on the DOE with different designed feature sizes, we also designed another three DOEs: 1500×1500 features with the feature size 0.0133 mm; 1000×1000 features with the feature size 0.02 mm and 500×500 features with the feature size 0.04 mm. They all perform very well under the totally coherent illumination as designed with the sharp edges. However under partially coherent illumination, the diffraction patterns will get blurred, similar to Fig. 6.6. Their diffraction patterns with different spatial coherence width are calculated, and their visibility are plotted in Fig. 6.7 (a). When the illuminating beam is totally incoherent (spatial coherence width approaches 0), we lost the visibility for all the DOEs, $v_m = 0$. As the spatial coherence width increases, the visibility of all the DOE increases as well. For the DOE with the smallest feature size 10 μm, the visibility of the diffraction pattern increases to 0.75, however the DOE with the biggest feature size 40 μm, we only got a visibility of 0.4. The visibility of the DOE diffraction pattern with small feature size increases much quickly than the DOE diffraction pattern with large feature size. For the same

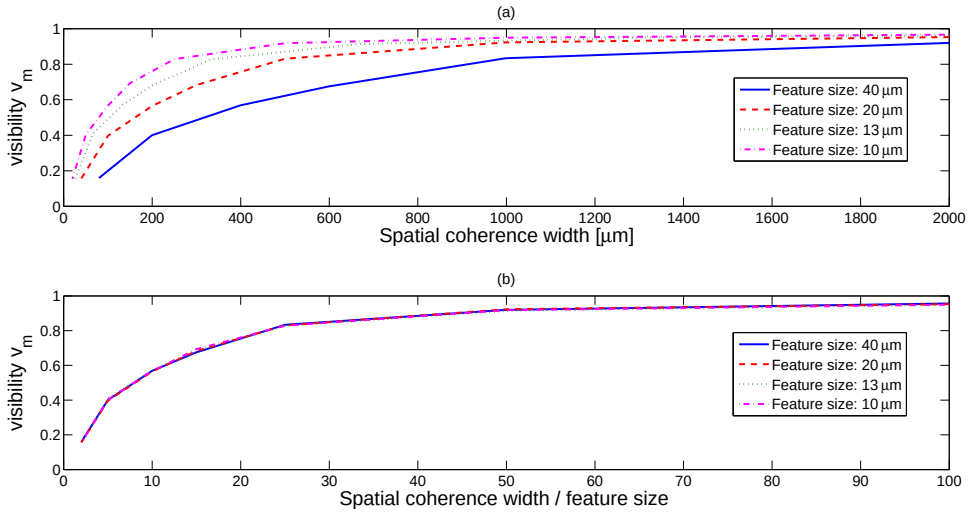


Fig. 6.7: (a): Visibility of the DOE diffraction pattern under spatially coherent illumination with different spatial coherence width. (b): Replot of (a) with the x axis of the ratio of coherence width and feature size.

partially coherent illumination, the DOE with the small feature size performs better than the DOE with the large feature size. If we replot Fig. 6.7 (a), changing the x axis to the ratio of the spatial coherence width to the feature size instead of spatial coherence width, the result is plotted in Fig. 6.7 (b). We found that all the curves are overlapped with each other, which means the ratio of the spatial coherence width to the feature size determines the performance of the DOE diffraction pattern under partially coherent illumination.

From the simulation results, we find the DOE can also work under partially coherent illumination, if the spatial coherence width of the light source is much bigger than the designed feature size of DOE. The DOE with small feature size is more adapted to the partially coherent beam, however it increases the manufacturing cost. By designing the DOEs, we could chose the appropriate feature size, to balance the cost and the efficiency of the DOEs.

6.5 Experiment results

In the last section, we have simulated the diffraction pattern of DOE under partially spatial coherent illumination. In this section, an experiment is designed and carried out as shown in Fig. 6.8 to demonstrate the theoretical model we have developed. A HeNe Laser is spatial filtered with a lens and a pinhole, then collimated into a plane wave, which is assumed to be fully spatial coherent. We first remove the diffuser and the lens L2, and thus have a coherent plane wave irradiating the DOE and the diffraction pattern is captured by a CCD sensor in an optical Fourier transform

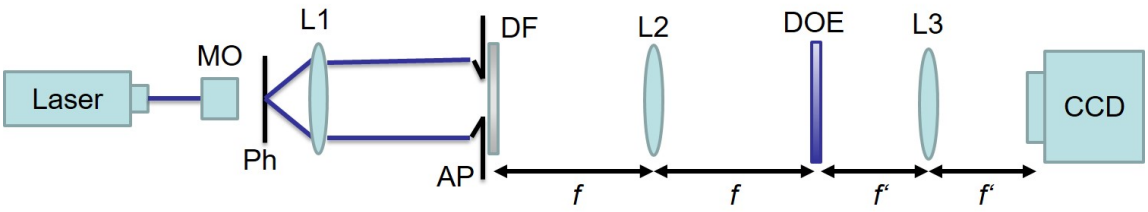


Fig. 6.8: Schematic of the experiment setup: L1,L2,L3, lens; MO: microscope objective; Ph, pinhole; DF, diffuser; AP: adjustable aperture.

system (The DOE diffraction profile is designed at the Fourier plane using the Gerchberg-Saxton algorithm introduced in Section 5.1, therefore we use a lens to measure its diffraction pattern at the Fourier plane). As the simulation results in the last section, we expect that if the light source is partially spatial coherent, the diffraction pattern captured by the CCD will get blurred, depending on the spatial coherence width. In order to generate a partially spatial coherent field, we could insert a diffuser and the lens L2 back into the light path between of collimated lens L1 and the DOE (Fig. 6.8), and rotate it [98, 99]. According to adjusting the aperture size before the diffuser, we can manipulate the spatial coherence width at the DOE plane, then measure the diffraction pattern of DOE. In our experiment, instead of rotating the diffuser to get a partially spatial coherent field, we use hundreds of stationary speckle field to equivalently simulate the partially spatial coherent field. First a spatial coherent speckle field is used to illuminate our DOE, and the diffraction pattern is recorded. We move the diffuser with every step 0.1 mm to generate a new spatial coherent speckle field, and record the diffraction pattern of this speckle illumination. 200 diffraction patterns of different speckle illumination are at last added together, their mean intensity is equivalent to captured intensity in the rotating diffuser case.

The diffraction pattern of DOE we used is designed as an owl. Because the size of the designed intensity profile is much bigger than the CCD sensor, only a part of the intensity profile is captured and shown in Fig. 6.9. Under totally coherent illumination, the diffraction pattern of this DOE is shown in Fig. 6.9 (a). A shape intensity profile can be seen as designed, however, the totally coherent illumination also introduces speckle near the profile. If the illumination is partially spatial coherent, the diffraction pattern will become blurred. In Fig. 6.9 (b) the aperture size before the DOE is 5 mm, the spatial coherence width is about 40 μm according to the Van Cittert-Zernike theorem. With such a big aperture size, the designed DOE intensity profile can barely be identified. If we reduce the aperture size, with 2 mm and 1 mm, the partially coherent fields with much larger spatial coherence width of 100 μm and 200 μm are generated. Their diffraction pattern are shown in Fig. 6.9 (c) and (d). As the spatial coherence width gets larger, the diffraction pattern goes more clear, and approaches to the designed profile. Note that the partially spatial coherent

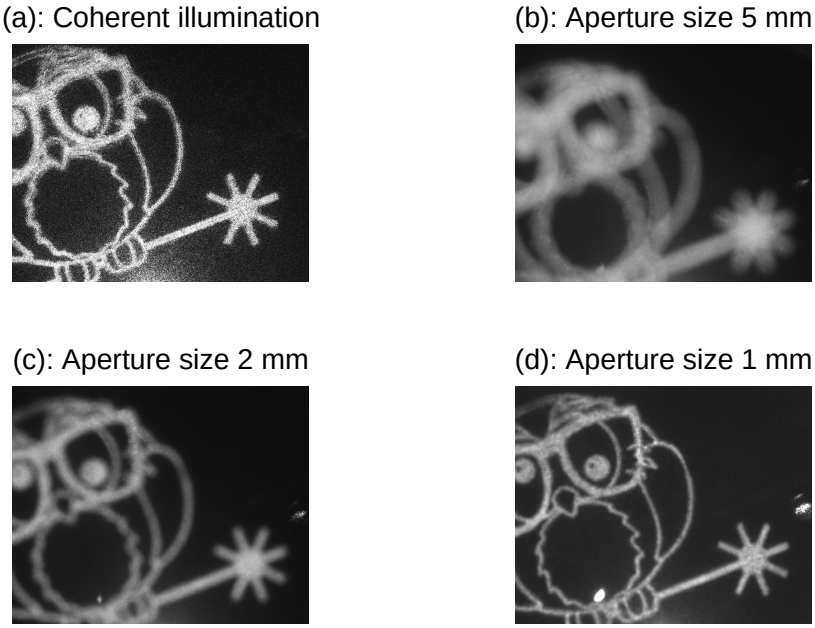


Fig. 6.9: DOE diffraction patterns under (a) totally coherent illumination, and partially spatial coherent illumination, the aperture size before the diffuser are (b) 5 mm, (c) 2 mm, (d) 1 mm.

illumination does not produce speckle in the diffraction plane, which makes partially spatial coherent illumination better than the laser in some imaging applications [100, 101].

6.6 Conclusion

In this chapter, we examined some aspects of coherence theory of the light, in particular light that is partially coherent. The partially coherent beam, can be consider to be the transitional status between the two extremes coherent and incoherent cases, is difficult to simulate. In order to describe the propagation of a partially coherent beam, we treated the beam as a summation of series spatially coherent random fields on an intensity basis over the integration time of the optical detection. After verify the proposed technique, the diffraction patterns of DOE under partially coherent illumination are investigated. Although the DOEs are designed for coherent illumination, we found that they also work under partially coherent illumination. And that the performance depends on the ratio of the coherence width to the designed feature size. The smaller the feature size of DOE, the better performs the DOE under partially coherent illumination. At the end of the chapter, we design a experiment,

which shows the DOEs do work under partially coherent illumination as simulated above.

7 Summary and outlook

This thesis addressed several problems in the digital optics. In modern optics, the photoelectric sensors like CCD / CMOS play very important role in recording optical signals. With the help of photoelectric sensors, the continuous optical signals are digitised, and are available for the subsequent processing.

The propagation of the light is one of the most important topic in the optics. In Chap 2 and 3, the Fresnel transform is detailed introduced and investigated. To calculate the Fresnel transform numerically, we uniformly sample the integration in Cartesian and cylindrical coordinate systems. It has been found that, the uniformly sampling will produce infinity replicas, which introduces error by insufficiently sampling of the input plane. The types of replicas are fundamental different depending on the sampling in different coordinate systems. As a 2D integration, it is very time costing to solve the Fresnel transform directly by sampling the input field. Two FFT based techniques are introduced, which greatly decrease the calculation time of the Fresnel transform. However due the uniformly sampling, the FFT based techniques also suffer the replica problems. To overcome this shortage, the analytical solutions of the Fresnel transform are investigated. We begin with the easy case of a perfect converging input field. According to comparing the four different methods, we found all the solutions are not closed form, they are written in infinity orders of polynomials. The third method with Zernike polynomials performs the best of all, it converges very quickly at nearly all the locations in the focal region. After that, we systematically studied the ENZ theory, and proposed the SENZ, which simplified the ENZ. At the end of Chap. 3, we use the same ideal in SENZ to calculate the Fresnel transform semi-analytically.

In Chap. 4, PSI as a wide used holographic technique is introduced. In our lens-less setup, we use Fresnel transform to reconstruct the object field, both intensity and phase. The resolution limits of the system are given. On one hand, it depends on the pixel size of the photoelectric sensors; on the other hand, it is limited by the numerical aperture D/z of the system. To approach the theoretical resolution limits, we analyzed different types of the noise in the system, and used correspondent techniques to reduce their effect of the system. After handling these noise sources, we have experimentally approached the theoretical resolution limits. At the end of Chap. 4, we used the PSI setup to measure the thickness of lens. Despite of the high

accuracy of the measurement, the range of measurement (surface slope) is limited by the resolution limits of the PSI system. For the future work, it will be meaningful, using different techniques to enlarge the range of measurement.

Phase retrieval as another technique to measure the phase information is discussed in Chap. 5. Due its simple setup (without reference wave) it was widely used in the optics. However as a iterative process, its convergence is not robust, by inappropriate initial phase estimating, the algorithm would converges very slowly, even converges to an incorrect result. Combining with the holographic setup, we proposed a simple way to measure the phase using iterative process. With the priori phase information, the phase can be calculated in few iterations for the off axis and speckle illumination cases.

In the last chapter, we investigated the partially coherent beam and its propagation. Instead of solving the 4D mutual coherence integral for the partially coherent beam, we used the speckle fields to simulate the partially coherent beam temporally. To verify the proposed simulation technique, we compare our simulation result to the analytical solution of the Thompson-Wolf experiment, which shows a high coincide between the two results. At the end, we simulate the diffraction pattern of DOE under partially coherent beam, and found the designed the ratio of the spatial coherence width to the feature size of DOE is an important factor to determine the performance of DOE under partially coherent illumination. However it is a prediction due simulation, it is important to experimentally measure the intensity profile of the DOE with different designed feature size in the future. This would provide us a way to beam shaping the partially coherent beam like LED using DOE by manipulating the spatial coherence width of the light source.

8 Publications

Journal publications

(1) Y. Wu; D. P. Kelly. Paraxial light distribution in the focal region of a lens: a comparison of several analytical solutions and a numerical result. Journal of Modern Optics, Vol. 61, pp. 57-67, 2014.
(2) Y. Wu; J. P. Ryle; S. Liu; D. P. Kelly; A. Stern. Experimental evaluation of inline free-space holography systems. Applied Optics, Vol. 54, pp. 3991-4000, 2015.
(3) Y. Wu; M. Hillenbrand; L. Zhao; S. Sinzinger; D. P. Kelly. The Fresnel transform as a projection onto a Nijboer-Zernike basis set. Optics Letters. Vol. 40, pp. 3472-3475, 2015.
(4) L. Zhao; Y. Wu; D. P. Kelly; J. T. Sheridan. Iterative reconstruction of digital holograms from three intensity measurements. Opt. Eng. Vol. 55, pp. 033106, 2016.

Conference contributions

(1) Y. Wu; D. P. Kelly. Monte-Carlo-Simulation von Licht im getrübten Medium. DGAO 2012.
(2) Y. Wu; D. P. Kelly. Paraxialen Lichtverteilung im Fokusbereich der Linse: ein Vergleich von mehreren analytischen Lösungen. DGAO 2013.
(3) Y. Wu; D. P. Kelly; A. Stern; S. Sinzinger. Fluctuations in the intensity read out of CCD/CMOS arrays in digital holographic setups: an experimental investigation. SPIE Sensing Technology Applications, 911708-911708-11, 2014
(4) Y. Wu; D. P. Kelly. Discrete numerical implementation of the Fresnel transform in Cartesian and cylindrical coordinate systems. DGAO 2014.
(5) L. Zhao; Y. Wu; C. Guo; D. P. Kelly; J. T. Sheridan. Reconstruction of digital holograms from three intensity measurements. SPIE 9599, Applications of Digital Image Processing XXXVIII 95992U, 2015.

References

[1] N. Streibl, K.-H. Brenner, A. Huang, J. Jahns, J. Jewell, A. W. Lohmann, D. A. B. Miller, M. Murdocca, M. E. Prise, and T. Sizer, "Digital optics," *Proceedings of the IEEE* **77**(12), pp. 1954–1969, 1989.

[2] F. Zernike, "Phase contrast, a new method for the microscopic observation of transparent objects," *Physica* **9**(7), pp. 686 – 698, 1942.

[3] P. Marquet, B. Rappaz, P. J. Magistretti, E. Cuche, Y. Emery, T. Colomb, and C. Depeursinge, "Digital holographic microscopy: a noninvasive contrast imaging technique allowing quantitative visualization of living cells with subwavelength axial accuracy," *Opt. Lett.* **30**, pp. 468–470, Mar 2005.

[4] D. P. Kelly, B. M. Hennelly, N. Pandey, T. J. Naughton, and W. T. Rhodes, "Resolution limits in practical digital holographic systems," *Optical Engineering* **48**(9), pp. 095801–095801–13, 2009.

[5] J. Goodman, *Introduction to Fourier optics*, Mc Graw Hill, 1996.

[6] R. S. Longhurst, *Geometrical and Physical Optics*, Orient BlackSwan, 1986.

[7] Y. Wu, J. P. Ryle, S. Liu, D. P. Kelly, and A. Stern, "Experimental evaluation of inline free-space holography systems," *Appl. Opt.* **54**, pp. 3991–4000, May 2015.

[8] A. W. Lohmann, *Optical Information Processing*, Universitätsverlag Ilmenau, 2006.

[9] A. Stern, "Sampling of linear canonical transformed signals," *Signal Processing* **86**, pp. 1421–1425, 2006.

[10] D. P. Kelly, "Numerical calculation of the fresnel transform," *J. Opt. Soc. Am. A* **31**, pp. 755–764, Apr 2014.

[11] A. D. Poularikas, *Handbook of Formulas and Tables for Signal Processing*, CRC Press, 1998.

[12] G. N. Watson, *A treatise on the theory of Bessel functions*, Ann Arbor, Michigan: University of Michigan Library, 2005.

[13] J. Fourier, *Theorie Analytique de la Chaleur*, Firmin Didot, 1822.

[14] M. Born and E. Wolf, *Principles of Optics*, Cambridge University Press, Oct. 1999.

[15] B. R. A. Nijboer, "The diffraction theory of optical aberration part i," *Physica X* **8**, pp. 679–692, 1943.

[16] B. R. A. Nijboer, "The diffraction theory of optical aberration part ii," *Physica XIII* **10**, pp. 605–620, 1947.

[17] Q. Cao, "Generalized jinc functions and their application to focusing and diffraction of circular apertures," *J. Opt. Soc. Am. A* **20**, pp. 661–667, Apr 2003.

[18] G. B. M. A. Gray and T. M. MacRobert, *A treatise on Bessel Functions*, London, Macmillan, 1922.

[19] D. P. Kelly, J. T. Sheridan, and W. T. Rhodes, "Finite-aperture effects for fourier transform systems with convergent illumination. part i: 2-d system analysis," *Optics Communications* **263**(2), pp. 171 – 179, 2006.

[20] D. P. Kelly, B. M. Hennelly, J. T. Sheridan, and W. T. Rhodes, "Finite-aperture effects for fourier transform systems with convergent illumination. part ii: 3-d system analysis," *Optics Communications* **263**(2), pp. 180 – 188, 2006.

[21] E. Lommel, "Die beugungserscheinungen einer kreisrunden oeffnung und eines kreisrunden schirmschens theoretisch und experimentell bearbeitet," *Abh. Bayer. Akad. Math. Naturwiss.* **15**, pp. 233–328, 1985.

[22] M. Abramowitz and I. A. Stegun, *Handbook of mathematical functions: with Formulas, Graphs, and Mathematical Tables*, Dover publications, 1964.

[23] B. R. A. Nijboer, *The diffraction theory of aberrations.* PhD thesis, Groningen University, 1942.

[24] J. L. Noyes, *Numerical Computation with Mathematica: Basic Analysis and Visualization*, Wittenberg University press, 2009.

[25] Y. Wu and D. P. Kelly, "Paraxial light distribution in the focal region of a lens: a comparison of several analytical solutions and a numerical result," *J. Mod. Opt.* **61**, pp. 57–67, 2014.

[26] A. J. E. M. Janssen, "Extended nijboer-zernike approach for the computation of optical point-spread functions," *J. Opt. Soc. Am. A* **19**, pp. 849–857, May 2002.

[27] J. J. M. Braat, P. Dirksen, and A. J. E. M. Janssen, "Assessment of an extended nijboer zernike approach for the computation of optical point-spread functions," *J. Opt. Soc. Am. A* **19**, pp. 858–870, 2002.

[28] A. J. E. M. Janssen, J. J. M. Braat, and P. Dirksen, "On the computation of the nijboer-zernike aberration integrals at arbitrary defocus," *J. Mod. Opt.* **51**, pp. 687–703, 2004.

[29] J. J. Braat, P. Dirksen, A. J. Janssen, S. van Haver, and A. S. van de Nes, "Extended nijboer zernike approach to aberration and birefringence retrieval in a high-numerical-aperture optical system," *J. Opt. Soc. Am. A* **22**, pp. 2635–2650, Dec 2005.

[30] S. van Haver, J. J. M. Braat, A. J. E. M. Janssen, O. T. A. Janssen, and S. F. Pereira, "Vectorial aerial-image computations of three-dimensional objects based on the extended nijboer-zernike theory," *J. Opt. Soc. Am. A* **26**, pp. 1221–1234, May 2009.

[31] S. van Haver, *The Extended Nijboer-Zernike Diffraction Theory and its Applications*. PhD thesis, Delft University of Technology, 2010.

[32] S. van Haver and A. J. E. M. Janssen, "Advanced analytic treatment and efficient computation of the diffraction integrals in the extended nijboer-zernike theory," *Journal of the European Optical Society - Rapid publications* **8**(0), 2013.

[33] S. van Haver and A. J. E. M. Janssen, "Truncation of the series expressions in the advanced enz-theory of diffraction integrals," *Journal of the European Optical Society - Rapid publications* **9**(0), 2014.

[34] F.Zernike, "Beugungstheorie des schneidenverfahrens und seiner verbesserten form, der phasenkontrastmethode," *Physica* **1**, pp. 689–704, 1934.

[35] R. J. Noll, "Zernike polynomials and atmospheric turbulence∗," *J. Opt. Soc. Am.* **66**, pp. 207–211, Mar 1976.

[36] B. Richards and E. Wolf, "Electromagnetic diffraction in optical systems. ii. structure of the image field in an aplanatic system," *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* **253**(1274), pp. 358–379, 1959.

[37] F. W. J. Olver, D. W. Lozier, R. F. Boisvert, and C. W. Clark, *NIST Handbook of Mathematical Functions*, Cambridge University Press, 2010.

[38] V. N. Mahajan and G. Dai, "Orthonormal polynomials in wavefront analysis: analytical solution," *J. Opt. Soc. Am. A* **24**, pp. 2994–3016, Sep 2007.

[39] B. H. Shakibaei and R. Paramesran, "Recursive formula to compute zernike radial polynomials," *Opt. Lett.* **38**, pp. 2487–2489, Jul 2013.

[40] J. Goodman, *Statistical Optics*, John Wiley & Sons, INC, 2000.

[41] D. Li, D. P. Kelly, and J. T. Sheridan, "Three-dimensional static speckle fields. part i. theory and numerical investigation," *J. Opt. Soc. Am. A* **28**, pp. 1896–1903, Sep 2011.

[42] D. Li, D. P. Kelly, and J. T. Sheridan, "Three-dimensional static speckle fields. part ii. experimental investigation," *J. Opt. Soc. Am. A* **28**, pp. 1904–1908, Sep 2011.

[43] Y. Wu, M. Hillenbrand, L. Zhao, S. Sinzinger, and D. P. Kelly, "Fresnel transform as a projection onto a nijboer zernike basis set," *Opt. Lett.* **40**, pp. 3472–3475, Aug 2015.

[44] E. N. Leith and J. Upatnieks, "Reconstructed wavefronts and communication theory," *J. Opt. Soc. Am.* **52**(10), pp. 1123–1128, 1962.

[45] I. Yamaguchi and T. Zhang, "Phase-shifting digital holography," *Opt. Lett.* **22**, pp. 1268–1270, Aug 1997.

[46] D. Kelly, J. Healy, B. Hennelly, and J. Sheridan, "Quantifying the 2.5d imaging performance of digital holographic systems," *Journal of the European Optical Society - Rapid publications* **6**(0), 2011.

[47] D. P. Kelly and D. Claus, "Filtering role of the sensor pixel in fourier and fresnel digital holography," *Appl. Opt.* **52**, pp. A336–A345, Jan 2013.

[48] D. P. Kelly, B. M. Hennelly, C. McElhinney, and T. J. Naughton, "A practical guide to digital holography and generalized sampling," *Optics and Photonics for Information Processing II* **7072**(1), p. 707215, SPIE, 2008.

[49] A. Stern and B. Javidi, "Sampling in the light of wigner distribution: errata," *J. Opt. Soc. Am. A* **21**(10), pp. 2038–2038, 2004.

[50] A. Stern and B. Javidi, "Shannon number and information capacity of three-dimensional integral imaging," *J. Opt. Soc. Am. A* **21**(9), pp. 1602–1612, 2004.

[51] A. Stern and B. Javidi, "Improved-resolution digital holography using the generalized sampling theorem for locally band-limited fields," *J. Opt. Soc. Am. A* **23**, pp. 1227–1235, May 2006.

[52] A. Stern and B. Javidi, "Analysis of practical sampling and reconstruction from fresnel fields," *Opt. Eng.* **43**, pp. 239–250, 2004.

[53] G. Å. Slettemoen and J. C. Wyant, "Maximal fraction of acceptable measurements in phase-shifting speckle interferometry: a theoretical study," *J. Opt. Soc. Am. A* **3**, pp. 210–214, Feb 1986.

[54] A. Stern and N. S. Kopeika, "Analytical method to calculate optical transfer functions for image motion and vibrations using moments," *J. Opt. Soc. Am. A* **14**, pp. 388–396, Feb 1997.

[55] A. Stern and N. S. Kopeika, "General restoration filter for vibrated-image restoration," *Appl. Opt.* **37**, pp. 7596–7603, Nov 1998.

[56] K. Kinnstaetter, A. W. Lohmann, J. Schwider, and N. Streibl, "Accuracy of phase shifting interferometry," *Appl. Opt.* **27**, pp. 5082–5089, Dec 1988.

[57] R. L. Powell and K. A. Stetson, "Interferometric vibration analysis by wavefront reconstruction," *J. Opt. Soc. Am.* **55**, pp. 1593–1598, Dec 1965.

[58] P. J. de Groot, "Vibration in phase-shifting interferometry," *J. Opt. Soc. Am. A* **12**, pp. 354–365, Feb 1995.

[59] N. Verrier, M. Gross, and M. Atlan, "Phase-resolved heterodyne holographic vibrometry with a strobe local oscillator," *Opt. Lett.* **38**, pp. 377–379, Feb 2013.

[60] A. V. Oppenheim, R. W. Schafer, and J. R. Buck, *Discrete time signal processing*, Prentice Hall, Inc., 2st ed., 1999.

[61] T. Colomb, F. Montfort, J. Kühn, N. Aspert, E. Cuche, A. Marian, F. Charrière, S. Bourquin, P. Marquet, and C. Depeursinge, "Numerical parametric lens for shifting, magnification, and complete aberration compensation in digital holographic microscopy," *J. Opt. Soc. Am. A* **23**, pp. 3177–3190, Dec 2006.

[62] T. Colomb, S. Krivec, H. Hutter, A. A. Akatay, N. Pavillon, F. Montfort, E. Cuche, J. Kühn, C. Depeursinge, and Y. Emery, "Digital holographic reflectometry," *Opt. Express* **18**, pp. 3719–3731, Feb 2010.

[63] M. Gross and M. Atlan, "Digital holography with ultimate sensitivity," *Opt. Lett.* **32**, pp. 909–911, Apr 2007.

[64] A. Sharma, G. Sheoran, Z. Jaffery, and Moinuddin, "Improvement of signal-to-noise ratio in digital holography using wavelet transform," *Optics and Lasers in Engineering* **46**(1), pp. 42 – 47, 2008.

[65] Y. Frauel, E. Tajahuerce, M.-A. Castro, and B. Javidi, "Distortion-tolerant three-dimensional object recognition with digital holography," *Appl. Opt.* **40**, pp. 3887–3893, Aug 2001.

[66] Y. Wu, D. P. Kelly, A. Stern, and S. Sinzinger, "Fluctuations in the intensity read out of ccd/cmos arrays in digital holographic setups: an experimental investigation," *Proceedings of SPIE* **9117**, SPIE, 2014.

[67] W. J. Dallas, I. Matsubara, and C. C. Yu, "High dynamic range holographic microscopy," in *Biomedical Optics and 3-D Imaging*, *Biomedical Optics and 3-D Imaging* , p. DTu1C.2, Optical Society of America, 2012.

[68] R. M. Groves, G. Pedrini, and W. Osten, "Real-time extended dynamic range imaging in shearography," *Appl. Opt.* **47**, pp. 5550–5556, Oct 2008.

[69] A. S. Y. Rivenson and B. Javidi, "Improved depth resolution by single-exposure in-line compressive holography," *Appl. Opt* **52**, pp. A223–A231, 2013.

[70] M. Lehmann, "Optimization of wavefield intensities in phase-shifting speckle interferometry," *Optics Communications* **118**(3–4), pp. 199 – 206, 1995.

[71] T. Maack, R. Kowarschik, and G. Notni, "Optimum lens aperture in phase-shifting speckle interferometric setups for maximum accuracy of phase measurement," *Appl. Opt.* **36**, pp. 6217–6224, Sep 1997.

[72] U. Grenander, *Probability and Statistics: The Harald Cramer Volume*, Wiley, 1st ed., 1959.

[73] I. Y. M. Y. G. M. Fucai Zhang, Jesus D. R. Valera, "Vibration analysis by phase shifting digital holography," *Optical Systems And Technologies* **11**(5), pp. 297–299, 2004.

[74] G. M. Pedrini G, Osten W, "High-speed digital holographic interferometry for vibration measurement," *Appl. Opt.* **45**(15), pp. 3456–3462, 2006.

[75] S. Schedin, G. Pedrini, H. J. Tiziani, and F. M. Santoyo, "Simultaneous three-dimensional dynamic deformation measurements with pulsed digital holography," *Appl. Opt.* **38**, pp. 7056–7062, Dec 1999.

[76] D. Claus, "High resolution digital holographic synthetic aperture applied to deformation measurement and extended depth of field method," *Appl. Opt.* **49**, pp. 3187–3198, Jun 2010.

[77] U. Schnars and W. Jueptner, *Digital Holography: Digital Hologram Recording, Numerical Reconstruction, and Related Techniques*, Springer, 1st ed., 1992.

[78] R. Cusack, J. M. Huntley, and H. T. Goldrein, "Improved noise-immune phase-unwrapping algorithm," *Appl. Opt.* **34**, pp. 781–789, Feb 1995.

[79] J. Meneses, T. Gharbi, and P. Humbert, "Phase-unwrapping algorithm for images with high noise content based on a local histogram," *Appl. Opt.* **44**, pp. 1207–1215, Mar 2005.

[80] M. A. Navarro, J. C. Estrada, M. Servin, J. A. Quiroga, and J. Vargas, "Fast two-dimensional simultaneous phase unwrapping and low-pass filtering," *Opt. Express* **20**, pp. 2556–2561, Jan 2012.

[81] R. W. Gerchberg and W. O. Saxton, "A practical algorithm for the determination of the phase from image and diffraction plane pictures," *Optik* **35**, pp. 237 – 246, 1972.

[82] J. R. Fienup, "Phase retrieval algorithms: a comparison," *Appl. Opt.* **21**, pp. 2758–2769, Aug 1982.

[83] M. J. Thomson, J. Liu, and M. R. Taghizadeh, "Iterative algorithm for the design of free-space diffractive optical elements for fiber coupling," *Appl. Opt.* **43**, pp. 1996–1999, Apr 2004.

[84] S. Vorndran, J. M. Russo, Y. Wu, S. A. Pelaez, and R. K. Kostuk, "Broadband gerchberg-saxton algorithm for freeform diffractive spectral filter design," *Opt. Express* **23**, pp. A1512–A1527, Nov 2015.

[85] A. J. Caley, M. J. Thomson, J. Liu, A. J. Waddie, and M. R. Taghizadeh, "Diffractive optical elements for high gain lasers with arbitrary output beam profiles," *Opt. Express* **15**, pp. 10699–10704, Aug 2007.

[86] N. C. Gallagher and B. Liu, "Method for computing kinoforms that reduces image reconstruction error," *Appl. Opt.* **12**, pp. 2328–2335, Oct 1973.

[87] A. Migukin, V. Katkovnik, and J. Astola, "Wave field reconstruction from multiple plane intensity-only data: augmented lagrangian algorithm," *J. Opt. Soc. Am. A* **28**, pp. 993–1002, Jun 2011.

[88] C. Guo, S. Liu, and J. T. Sheridan, "Iterative phase retrieval algorithms. i: optimization," *Appl. Opt.* **54**, pp. 4698–4708, May 2015.

[89] T. M. L. Megel, D. P. Kelly and S. Sinzinger, "Iterative phase retrieval and the important role played by initial conditions," *Fringe 2013 7th International Workshop on Advanced Optical Imaging and Metrology* , Springer Berlin Heidelberg, 2014.

[90] Y.-Y. Cheng and J. C. Wyant, "Phase shifter calibration in phase-shifting interferometry," *Appl. Opt.* **24**, pp. 3049–3052, Sep 1985.

[91] G.-S. Han and S.-W. Kim, "Numerical correction of reference phases in phase-shifting interferometry by iterative least-squares fitting," *Appl. Opt.* **33**, pp. 7321–7325, Nov 1994.

[92] L. Zhao, Y. Wu, D. P. Kelly, and J. T. Sheridan, "Iterative reconstruction of digital holograms from three intensity measurements," *Optical Engineering* **55**(3), p. 033106, 2016.

[93] G. Liu and P. D. Scott, "Phase retrieval and twin-image elimination for in-line fresnel holograms," *J. Opt. Soc. Am. A* **4**, pp. 159–165, Jan 1987.

[94] J. R. Fienup, "Reconstruction of a complex-valued object from the modulus of its fourier transform using a support constraint," *J. Opt. Soc. Am. A* **4**, pp. 118–123, Jan 1987.

[95] W. Martienssen and E. Spiller, "Coherence and fluctuations in light beams," *American Journal of Physics* **32**, pp. 919–926, 1964.

[96] H. Arsenault and S. Lowenthal, "Partial coherence in the image of an object illuminated with laser light through a moving diffuser," *Optics Communications* **1**, pp. 451–453, Apr. 1970.

[97] B. J. Thompson and E. Wolf, "Two-beam interference with partially coherent light," *J. Opt. Soc. Am.* **47**, pp. 895–902, Oct 1957.

[98] G. Li, Y. Qiu, and H. Li, "Coherence theory of a laser beam passing through a moving diffuser," *Opt. Express* **21**, pp. 13032–13039, Jun 2013.

[99] J. Lehtolahti, M. Kuittinen, J. Turunen, and J. Tervo, "Coherence modulation by deterministic rotating diffusers," *Opt. Express* **23**, pp. 10453–10466, Apr 2015.

[100] J. A. Rodrigo and T. Alieva, "Illumination coherence engineering and quantitative phase imaging," *Opt. Lett.* **39**, pp. 5634–5637, Oct 2014.

[101] J. A. Rodrigo and T. Alieva, "Rapid quantitative phase imaging for partially coherent light microscopy," *Opt. Express* **22**, pp. 13472–13483, Jun 2014.