Computational Reconstruction Of Images From Optical Holograms

by

Weichang Li

Submitted to the Department of Ocean Engineering and
Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degrees of

Master of Science in Ocean Engineering

and

Master of Science in Electrical Engineering and Computer Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2002

Author .....
Department of Ocean Engineering
September 1, 2001

Certified by .....
Jerome H. Milgram
W. I. Koch Professor of Marine Technology
Thesis Supervisor

Accepted by .....
Henrik Schmidt
Chairman, Department Committee on Graduate Students

Read by .....
Arthur B. Baggeroer
Ford Professor of Engineering
Thesis Reader

Accepted by .....
Arthur C. Smith
Chairman, Department Committee on Graduate Students

Computational Reconstruction Of Images From Optical Holograms

by

Weichang Li

Submitted to the Department of Ocean Engineering
and
Department of Electrical Engineering and Computer Science
on September 1, 2001, in partial fulfillment of the
requirements for the degrees of
Master of Science in Ocean Engineering
and
Master of Science in Electrical Engineering and Computer Science

Abstract

This thesis describes the analysis, development and implementation of computational reconstruction of optical holograms, and its applications in marine biology and Holographic Particle Image Velocimetry (HPIV). Computational reconstruction of holograms is a process that recovers three-dimensional information from its holographic form computationally rather than optically. In this thesis, the reconstructions are represented by a sequence of plane images at programmable distances parallel to the associated hologram. The primary goal is to increase the reconstruction speed and improve the image quality, and ultimately lead to a computational platform that is efficient and robust for oceanographic applications such as for marine biology and HPIV. A signal processing path has been followed throughout this thesis. Reconstruction algorithms have been developed in wavenumber space and accelerated by the use of FFT algorithms. Modifications to the reconstruction kernel both in spatial domain and frequency domain have been incorporated to improve the reconstruction quality. The behavior of reconstruction process is described qualitatively. A number of focus measures are proposed for quantitative focus analysis based on a circular disc object. Optimal sampling of holograms is also discussed and found to be distance-dependent. A PC based computational reconstruction system has been constructed which in addition consists of a commercial memory-enhanced DSP engine and a high-resolution scanner. As compared to optical reconstruction, by using this system the reconstruction time of one hologram has been reduced by a factor of 100 to 383.3 minutes for recovering one $24.8mm * 24.8mm * 250mm$ volume at a resolution of $315pixel/mm$. Both simulation results of HPIV and experimental results of marine micro-organism are presented. A computationally efficient method has been devised for object-counting based on hologram images.

Thesis Supervisor: Jerome H. Milgram
Title: W. I. Koch Professor of Marine Technology

Thesis Reader: Arthur B. Baggeroer
Title: Ford Professor of Engineering

Acknowledgments

I am more than grateful to my advisor, Prof. Jerry Milgram, for taking the gamble and giving me the chance to work on this computational holography project, providing funds and guiding me through. I also wish to thank Prof. Art Baggeroer, for being my thesis reader on EECS part and providing insightful suggestions.

Lots of other people had provided help during the course of this thesis and deserve credits here. Thanks to Dr. Ed. Malkiel and Prof. Joseph. Katz of Johns Hopkins Univ. for providing all the optical holograms that have been used in this thesis. The discussion with Dr. Ed. Malkiel and Omar Alquaddoomi triggered the object counting part in this thesis. Also I would like to thank Dr. Edith Widder for sending me bioluminescence images, and Dr. Richard Lampitt for providing me the real corpepod image.

Thanks to Rick Rikoski for proofreading chapter 4 and 5, I know it is not easy.

Finally, I would like to thank my wife, Hua, for her unfailing support, both moral and practical, throughout the whole course.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Overview

This thesis describes the analysis, development and implementation of computational reconstruction of optical holograms, and its applications in marine biology and Holographic Particle Image Velocimetry (HPIV). Computational reconstruction of holograms is a process that recovers the three-dimensional information encoded in a hologram computationally rather than optically. In this thesis, the reconstructed results are represented by a sequence of plane images at programmable distances parallel to the associated hologram. The primary goal is to increase the reconstruction speed and improve the image quality, and ultimately to provide a computational platform that is efficient and robust for oceanographic applications such as for marine biology and HPIV. A signal processing path has been followed throughout this thesis. Reconstruction algorithms have been developed in wavenumber space and accelerated by the use of FFT algorithms. Modified kernel functions by spatial windowing and filtering have been incorporated to improve the reconstruction performance. The behavior of the reconstruction process is described as a matched filtering. A number of focus measures are proposed for focus analysis. Optimal sampling of holograms is also discussed and found to be distance-dependent. A PC based computational reconstruction system has been constructed which in addition consists of a commercial memory-enhanced DSP engine and a high-resolution scanner. Both simulation results of HPIV and experimental results of marine micro-organisms are presented. A computationally efficient method has been devised for object-counting based on hologram images.

1.2 Holography, Computational Holography

The invention of optical holography was a result of an "exercise in serendipity", as Dr. Dennis Gabor explained in his autobiography. He proposed the principle of wavefront reconstruction in [6], followed by two more lengthy papers [7][8], mainly considering the possible applications of holography to microscopy. The technique attracted mild interest until in 1960s when the concepts and technology were dramatically improved, mainly due to the invention of the laser and the off-axis holography technique introduced by E. N. Leith and J. Upatnieks [4]. Its applicability and practicality were vastly extended, which leads to successes in a variety of applications. The quality and realism of three dimensional images obtained by holography led to the development of a great interest in the field.

Computational holography is about computational synthesis of holograms for later display, or computational reconstruction of optical holograms. Computational synthesis of holograms began as early as 1964; researchers started to consider the computation, transmission and use of synthesized holograms to create images [23]. Early development of computational analysis of particle holograms [25] were often considered under far-field assumption which reduces the reconstruction to a modified Fourier transform. Barriers to get high-fidelity reconstruction computationally include the substantial computation capacity required to manipulate the tremendous amount of information embedded in a hologram, and the distortion and intrinsic noise inherent in the processes as well. In 1980's Onural and Scott [19] reported a method for the digital decoding of the in-line hologram, which is equivalent to a high order approximation that can reduce the twin-image effect. Unfortunately as the order increases, the resulted image quality is degraded by other artifacts. Computational reconstruction mainly involves the discretization of holograms, 2D Fourier transform, inverse transform and filtering. Most previous work is focused on computational implementation of optical processes, which is a good point to start, but not necessary the best solution for computational reconstruction. Instead of being viewed merely as computational implementation of optical holographic processes, the problem can also be considered as a inverse problem to which appropriate signal processing operations can be applied.

1.3 Motivations for Computational Reconstruction of Optical Holograms

Holographic particle image velocimetry (HPIV) and underwater hologrammetry constitute the two major application fields of computational reconstruction of holograms. Particle image velocimetry(PIV) measures fluid flow field by measuring the velocity of particles in a fluid. HPIV is capable of doing 3D velocity measurement in one shot, hence is ideal for measuring complex high-speed flow such as turbulent flow. The other application of underwater hologrammetry has been mainly used by marine biologists to find out the individual morphology, population pattern and mutual dependence of marine micro-organisms such as plankton. These are well known to be an essential part of the food chain. As an noninvasive, in-situ and 3D imaging technique, holography has become a promising tool for marine biologists, capable to provide the data necessary for the mathematical and conceptual modeling of the underwater ecosystem.

For all these and other applications, most of the present technology is photographic recording media. However, the real challenge lies in processing the huge amount of information contained in these holograms. Typically, it would take several days to reconstruct an object scene from one hologram using optical replay and high resolution digitizing systems. The laborious process can be unstable and inflexible. The image quality is influenced by optical abberations and limited by the replay setup. By computational reconstruction, the motivation is to develop alternative methodology and implementation systems that work computationally rather than optically, which are able to generate high fidelity reconstruction in a short amount of time. The long term hope is the ultimate development of a film free digital recording and reconstruction system which, as we believe, will be more efficient and robust than current optical systems, and easier to operate.

Chapter 2

Modeling and Analysis of Optical Holography

The principle and theory of optical holography have been well established since Dr. Gabor invented optical holography and published his pioneering work in [6][7][8], and has been vastly exploited in a broad range of applications after the invention of off-axis holography by E.N. Leith and J. Upatnieks [4]. Linear system treatment of scalar diffraction theory has also been well developed and can be found in Goodman's classical textbook on Fourier optics [10]. Therefore in this chapter we are not intending to present any new theory about optical holography. Rather, effort is made to formulate optical holography in a linear system frame and analyze it from signal processing perspective. A general formulation of both in-line holography and off-axis holography is proposed explicitly as a 2D linear convolution. This underlies the basis for computational reconstruction that will be discussed in subsequent chapters.

Optical holography, consisting of two steps including hologram recording and image reconstruction, is essentially a synthesis of interferometry and diffraction. A hologram as recorded, regardless of its media form, is a set of fringe patterns that results from the interference between an object wave and a reference wave. When being reconstructed, it manifests itself as a set of diffraction gratings. The associated diffracted field constitutes a reconstruction of the original object field. Generally in a typical holography setup, the object waves striking the hologram plane are the scattered waves from certain object sources. The recording of an optical field on holographic media such as film and plates is an intensity

mapping process. Therefore an optical hologram is an "energy detector" for the total interaction field. (Note that for long-wave cases such as acoustic holography and electronic holography, sensors are readily available for direct recording of a full wave field. Thus the holography technique of introducing a reference wave is not necessarily a startling achievement for long-wave cases.)

Both holographic recording and reconstruction are composed of a similar set of physical phenomena, namely interference and diffraction. Or as put by E. N. Leith and J. Upatnieks [4] in communications terminology, "a modulation, a frequency dispersion and a square-law detection". "Square-law detector" refers to the intensity mapping of a hologram recording. In addition to two linear data-containing terms, the intensity mapping also generates a DC component and nonlinear terms which are not linear operations. The DC component corresponds to transmitted reference light and can be eliminated in computational reconstruction. The nonlinear term is the object distribution intensity which is usually very weak and negligible if the reference wave is made much stronger than the object distribution. Under these assumptions, the interference becomes bipolar and linear treatment is justified.

According to Fourier optics [10], the propagation of the optical field from an object space onto the hologram plane, or the diffraction from an object distribution to the hologram plane, is essentially a linear all-pass filtering process with a quadratic phase response under the paraxial approximation. It has also often called the Fresnel transform [11]. The Fresnel transform characterizes the propagation of an optical field in the Fresnel region which is a continuum transition field having mixed properties of both the object space and the Fourier space. Under far field assumption, the Fresnel diffraction degenerates into the Fraunhofer diffraction. Due to the simplicity of Fraunhofer diffraction, most applications assume the far-field criteria being satisfied [26], such as particle field holography and HPIV. Issues such as transverse and axial resolutions have usually been exploited under the same simplification [2]. However, the same issues in Fresnel region have rarely been discussed explicitly. The reason is because when the near field is being considered, the analysis becomes much more complicated. In this sense classical Fourier optics theory is somewhat far-field oriented.

This chapter is divided into two parts: formulation and analysis. Formulation is carried out in section2.1 for simple disc objects. It proceeds by treating holography as a combination of bipolar interference (section 2.1.1) and diffraction (section 2.1.2). Ultimately the

general form is given in section 2.1.3, with extension to three dimensional object distribution discussed in section 2.1.4. In section 2.2, the analysis part is discussed. Based on Fourier optics and its physical realm, namely the plane wave decomposition and angular spectrum propagation, the finite hologram aperture effect is described. In addition an analogy of optical holography as a chirp carrier modulation has been established.

2.1 Linear System Modeling

In this section, we formulate both in-line holography and off-axis holography in a general mathematical representation. In the following discussion, the recording and reconstruction are covered simultaneously as they fundamentally embody the same set of physical processes.

2.1.1 Principle of Holography: Bipolar Interference

Holography is essentially an interference phenomenon, or a modulation process. In this section holographic recording and reconstruction are described mainly as wavefront encoding and decoding processes, for an arbitrary object distribution. The connection between the object wavefront and the associated object distribution will be explained in section 2.1.2.

Figures 2-1, 2-2, 2-3 and 2-4 show sketches of making holograms and reconstructing their optical fields for both in-line and off-axis holograms.

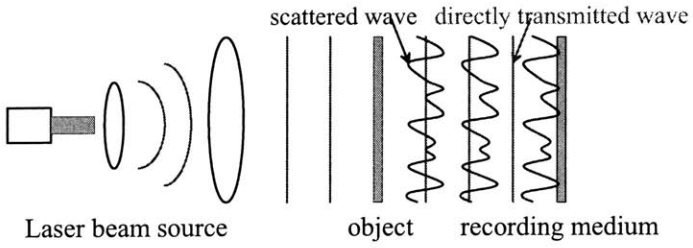

Figure 2-1: Optical system used to record an in-line hologram;

Wavefront Recording

For both in-line and off-axis holograms, the optical field that falls upon the holographic film is the sum of a reference beam, $r(x_h, y_h)$, and the object beam, $o(x_h, y_h)$. x_h and y_h are the

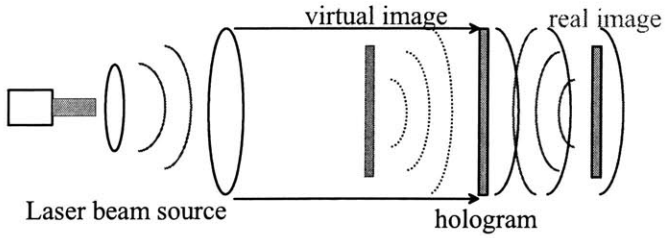

Figure 2-2: Optical system used to reconstruct the image from an in-line hologram;

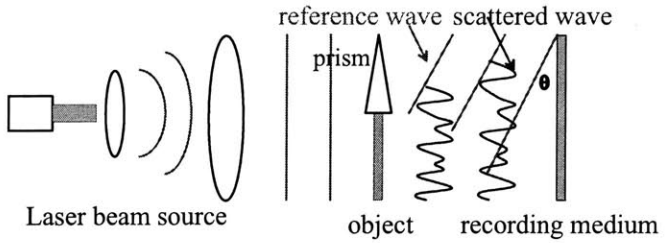

Figure 2-3: Hologram recording with an off-axis reference beam;

coordinates on the holographic film with the origin on the optical axis. The light intensity, I, which is what the film records, is the magnitude squared of the field.

$$\begin{aligned} I(x_h, y_h) &= |r(x_h, y_h) + o(x_h, y_h)|^2 \\ &= |r(x_h, y_h)|^2 + |o(x_h, y_h)|^2 + r(x_h, y_h)o^*(x_h, y_h) + r^*(x_h, y_h)o(x_h, y_h) \quad (2.1) \end{aligned}$$

* indicates complex conjugate.

Of the four terms on the right hand side of equation (2.1), the first one is simply a constant, which has no consequence. The second term is an out of focus representation of the image, usually called the halo. The third and fourth terms are the optical fields of the real and virtual images multiplied by a constant for an in-line hologram and by a spatial sinusoid for an off-axis hologram. It is these last two terms that contain the optical field information we wish to capture and record. Generally, these terms are made much larger than the undesirable second term by making r much larger than o. For in-line holograms, we assume this has been done and ignore the first two terms. For off-axis holograms, we presume the off-axis angle is large enough to spatially separate the images from the halo

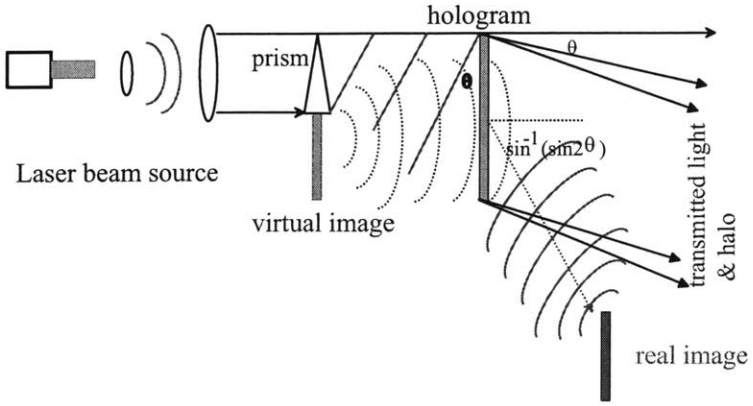

Figure 2-4: Image reconstruction from an off-axis hologram;

and again the second term can be ignored. Therefore a bipolar form of $I(x_h, y_h)$ becomes:

$$I(x_h, y_h) \quad = r(x_h, y_h)o^*(x_h, y_h) + r^*(x_h, y_h)o(x_h, y_h) \tag{2.2}$$

The description is given here in terms of the scalar theory of light (c.f. Goodman, 1996) [10]. An implicit aspect of the theory used here is that the functions refer to components of the reference and object beams that have the same polarization. For the representation of light waves, a time function of the form $\exp(-i\omega t)$ is implied and not written. The form of the reference wave is:

$$r(x_h, y_h) = A\exp(jkx_h \sin\theta) \tag{2.3}$$

where θ is the off-axis angle with respect to the x_h axis, A is the uniform real amplitude. The reference beam is presumed to be perpendicular to the y_h axis. For an in-line hologram, $\theta = 0$. k is the circular wave number of the laser light beam.

If we assume [10]:

1 The variations of exposure in the interference pattern remain within a linear region of the t_A vs E curve of the recording media.

2 The Modulation Transfer Function (MTF) of the recording material extends to sufficiently high spatial frequencies to record all the incident spatial structure.

3 The intensity A^2 of the reference wave is uniform across the recording material.

Then after dropping off any uniform bias transmittance, the amplitude transmittance of the developed film or plate can be written as:

$$t_h(x_h, y_h) = 2\beta A \mathbf{Re}\{\exp(ikx_h \sin\theta) o^*(x_h, y_h)\} \tag{2.4}$$

Where β is a constant related to the exposure time and the slope of the t_A vs E curve.

Wavefront Reconstruction

Suppose that the developed transparency is illuminated by an reference light identical to the one used for recording. Then the light transmitted by the transparency becomes:

$$\begin{aligned} p_h(x_h, y_h) &= r(x_h, y_h) t_h(x_h, y_h) \\ &= \beta A^2 \exp(j2kx_h \sin\theta) o^*(x_h, y_h) + \beta A^2 o(x_h, y_h) \end{aligned} \tag{2.5}$$

The first term is proportional to the conjugate of original object wave. It is deflected away from the optical axis at angle $sin^{-1}(2sin\theta)$ and forms a real image field. The second term is the original object wave multiplied by a constant factor. It propagates along the optical axis and generates a virtual image field. Since $\theta = 0$ for in-line holography, all components are overlapped in space consequently. In off-axis case, it is possible to separate all these terms in space by increasing θ.

2.1.2 Diffraction and The Fresnel Transformation

In transmission holography, one common physical phenomena, diffraction, is shared by both the recording and the reconstruction processes. During recording, part of the reference wave is scattered by a source distribution in the object space and the scattered light reaches the recording media, interfering with the directly transmitted portion of the reference beam and being recorded. In the reconstruction phase, a reference light is incident on the hologram which is now equivalent to a diffraction grating, then the resulting diffraction field constitutes a recovery of the original wavefield. In this section, we formulate the diffraction field of a given 2D object distribution function as a linear convolution output. Without a paraxial approximation, the convolution kernel turns out to be as an exact kernel as given below. Under the paraxial approximation, the exact kernel degenerates into the Fresnel

transform kernel which is a pure quadratic phase factor.

Here we use the Huygens-Fresnel theory of diffraction as predicted by the first Rayleigh-Sommerfeld solution [10]. When illuminated by a normally incident uniform plane wave, the presence of two-dimensional object sources $t_o(x_o, y_o)$ across a plane Σ modifies the amplitude at an observation point P in the diffraction field to:

$$o(P) = \frac{1}{j\lambda} \iint_{\Sigma} t_o(x_o, y_o) \frac{exp(jkr)}{r} \; cos\theta \; dx_o dy_o \tag{2.6}$$

Where λ is the laser wavelength and $k = 2\pi/\lambda$. Σ has coordinates (x_o, y_o, z) with z fixed. $z = 0$ at point P. r is the distance from point (x_o, y_o, z) to point P. $cos\theta$ is the so called obliquity factor [10]. If we assume that P is a point on hologram plane with coordinates $(x_h, y_h, 0)$, then r and the obliquity factor can be written as:

$$\begin{aligned} r &= \sqrt{(x_h - x_o)^2 + (y_h - y_o)^2 + z^2} \\ cos\theta &= \frac{z}{r} \end{aligned} \tag{2.7}$$

Therefore the object wave coming from the source and falling on the hologram becomes:

$$o(x_h, y_h) = \frac{z}{j\lambda} \iint_{\Sigma} t_o(x_o, y_o) \frac{exp(jk\sqrt{(x_h - x_o)^2 + (y_h - y_o)^2 + z^2})}{(x_h - x_o)^2 + (y_h - y_o)^2) + z^2} dx_o dy_o \tag{2.8}$$

The above equation has the form of a 2D linear convolution:

$$\begin{aligned} o(x_h, y_h) &= \iint_{\Sigma} t_o(x_o, y_o) h_e(x_h - x_o, y_h - y_o; z) dx_o dy_o \\ &= \{t(x, y) * * h_e(x, y; z)\} \mid_{x=x_h, y=y_h} \end{aligned} \tag{2.9}$$

where $**$ means 2D linear convolution, and

$$h_e(x, y; z) = \frac{z}{j\lambda} \frac{exp(jk\sqrt{x^2 + y^2 + z^2})}{x^2 + y^2 + z^2} \tag{2.10}$$

Since the representation in equation (2.10) is derived exactly without approximation, we call $h_e(x, y; z)$ as the exact kernel. It can be simplified using the paraxial approximation. If the lateral quantities $(x_o - x_h)$ and $(y_o - y_h)$ are small in comparison to the axial distance z, both the denominators and the arguments of the exponentials in equation (2.8) can be

approximated on this basis. The arguments of the exponentials must be carried out to one higher order in $(x_o - x_h)$ and $(y_o - y_h)$ to retain the crucial phase information. The resulting expression for the field on the hologram, o_p is called the paraxial approximation and is given by:

$$o_p(x_h, y_h) = \frac{exp(jkz)}{j\lambda z} \iint_{\Sigma} t_o(x_o, y_o) exp(\frac{jk}{2z}[(x_h - x_o)^2 + (y_h - y_o)^2]) dx_o dy_o \quad (2.11)$$

If we let

$$h_f(x, y; z) = \frac{1}{j\lambda z} exp[\frac{jk}{2z}(x^2 + y^2)] = \frac{1}{j\lambda z} exp[\frac{j\pi}{\lambda z}(x^2 + y^2)] \quad (2.12)$$

Then

$$\begin{aligned} o_p(x_h, y_h) &= \iint_{\Sigma} t_o(x_o, y_o) h_f(x_h - x_o, y_h - y_o; z) dx_o dy_o \\ &= \{t_o(x, y) * * h_f(x, y; z)\} \mid_{x=x_h, y=y_h} \end{aligned} \quad (2.13)$$

where the constant phase factor $exp(jkz)$ has been dropped off since it has no consequence to reconstructed images.

$h_f(x, y; z)$ is a 2D chirp function and has a quadratic phase structure. It is usually called the Fresnel transform kernel. One nice property of the Fresnel kernel is that its Fourier transform is analytically available and is also a 2D chirp function in the spatial frequency domain.

$$\begin{aligned} H_f(f_x, f_y; z) &= \mathcal{F}\{h_f(x, y; z)\} \\ &= \exp\left[-j\pi\lambda z(f_x^2 + f_y^2)\right] \end{aligned} \quad (2.14)$$

Corresponding to the concept of instantaneous frequency of a time signal, we define the local spatial frequency here as the derivative of the phase:

$$\begin{aligned} f_x(x) &= \frac{1}{2\pi} \frac{\partial(\frac{\pi}{\lambda z}(x^2 + y^2))}{\partial x} = \frac{x}{\lambda z} \\ f_y(y) &= \frac{1}{2\pi} \frac{\partial(\frac{\pi}{\lambda z}(x^2 + y^2))}{\partial y} = \frac{y}{\lambda z} \end{aligned} \quad (2.15)$$

We can further define the chirp rate as the derivative of the local frequency regarding the associated coordinate. Basically the chirp rate is the rate at which the frequency changes

with spatial variables and it is the same for both x and y directions in this case. By definition, its value is $C_r = 1/(\lambda z)$. Note that $\sqrt{\lambda z}$ is the characteristic length of diffraction in optics. We can notice that the Fresnel kernel has a linearly modulated local spatial frequency whose modulating rate is C_r. Obviously C_r decreases as z increases for fixed value of λ.

It is easy to see that the Fresnel kernel function is neither spatially limited nor band limited. From equation (2.15), we can find that if it is spatially truncated, the maximum local frequency of $h_f(x, y; z)$ is dependent on the truncating window size, and also increases as the value of z decreases. This property will be used later to determine the sampling rate and spatial extent of the discrete kernel function in section 3.1.

A number of other properties of the Fresnel kernel that worth to be mentioned here, which are very helpful for the intuitive understanding of the diffraction process.

First, the form of $H_f(f_x, f_y; z)$ implies that the Fresnel kernel is cascadable, i.e:

$$\begin{aligned} H_f(f_x, f_y; z1)H_f(f_x, f_y; z2) &= H_f(f_x, f_y; z1 + z2) \\ h_f(x, y; z1) * * h_f(x, y; z2) &= h_f(x, y; z1 + z2) \end{aligned} \tag{2.16}$$

Consequently,

$$\begin{aligned} H_f(f_x, f_y; z)H_f(f_x, f_y; -z) &= H_f(f_x, f_y; 0) = 1 \\ h_f(x, y; z) * * h_f(x, y; -z) &= \delta(x, y) \end{aligned} \tag{2.17}$$

Also $h_f^*(x, y; z) = h_f(x, y; -z)$.

Second, the convolution of $h_f(x, y; z)$ with a uniform plane wave $A\exp[j(kxsin\theta + kysin\beta)]$ whose directional cosine is $(sin\theta, sin\beta)$ becomes

$$\begin{aligned} & H_f(f_x, f_y; z)\mathcal{F}\{A\exp[j(kxsin\theta + kysin\beta)]\} \\ & = H_f(f_x, f_y; z)A\delta(f_x - \tfrac{sin\theta}{\lambda}, f_y - \tfrac{sin\beta}{\lambda}) \\ & = \exp\left[-j\pi\lambda z([\tfrac{sin\theta}{\lambda}]^2 + [\tfrac{sin\beta}{\lambda}]^2)\right] A\delta(f_x - \tfrac{sin\theta}{\lambda}, f_y - \tfrac{sin\beta}{\lambda}) \end{aligned} \tag{2.18}$$

Hence we can see that the effect is merely a constant phase change which is dependent on

the directional cosine:

$$\begin{aligned} & h_f(x,y;z) * *A\exp[j(kxsin\theta + kysin\beta)] \\ &= A\exp[j(kxsin\theta + kysin\beta)]\exp\left[-j\pi\lambda z([\tfrac{sin\theta}{\lambda}]^2 + [\tfrac{sin\beta}{\lambda}]^2)\right] \end{aligned} \tag{2.19}$$

Furthermore, the convolution of a constant A with $h_f(x,y;z)$ gives A itself.

$$\begin{aligned} H_f(f_x,f_y;z)A\delta(f_x,f_y) &= A\delta(f_x,f_y) \\ h_f(x,y;z) * *A &= A \end{aligned} \tag{2.20}$$

This intuitively makes sense because the free-space propagation of a uniform plane wave should change nothing to the wavefront.

2.1.3 Formulation of Holographic Recording and Reconstruction

Now we can combine the results of both section 2.1.1 and section 2.1.2 to generate a mathematical formulation for the whole process. Again we start by considering a two-dimensional object distribution $t_o(x_o,y_o;z_o)$ which is located at $z = z_o$ and parallel to the hologram plane. The expression of the associated hologram can be derived by substituting equation (2.8) or (2.13) into equation (2.4) :

$$\begin{aligned} t_h(x_h,y_h;z_o) &= \{\beta A\exp(jkx\sin\theta)[t_o^*(x,y;z_o) * *h^*(x,y;z_o)] \\ &\quad +\beta A\exp(-jkx\sin\theta)[t_o(x,y;z_o) * *h(x,y;z_o)]\}\ |_{x=x_h,y=y_h} \end{aligned} \tag{2.21}$$

Where $h(x,y;z_o) = h_e(x,y;z_o)$ without the paraxial approximation, otherwise $h(x,y;z_o) = h_f(x,y;z_o)$.

Similarly by substituting equation (2.8) or (2.13) into equation (2.5), we can express the light transmitted through the hologram (at $z = 0$) as:

$$\begin{aligned} p_h(x_h,y_h;z_o) &= \{\beta A^2\exp(j2kx\sin\theta)[t_o^*(x,y;z_o) * *h^*(x,y;z_o)] \\ &\quad +\beta A^2[t_o(x,y;z_o) * *h(x,y;z_o)]\}\ |_{x=x_h,y=y_h} \end{aligned} \tag{2.22}$$

Where again $h(x,y;z_o) = h_e(x,y;z_o)$ without paraxial approximation, and $h(x,y;z_o) = h_f(x,y;z_o)$ otherwise.

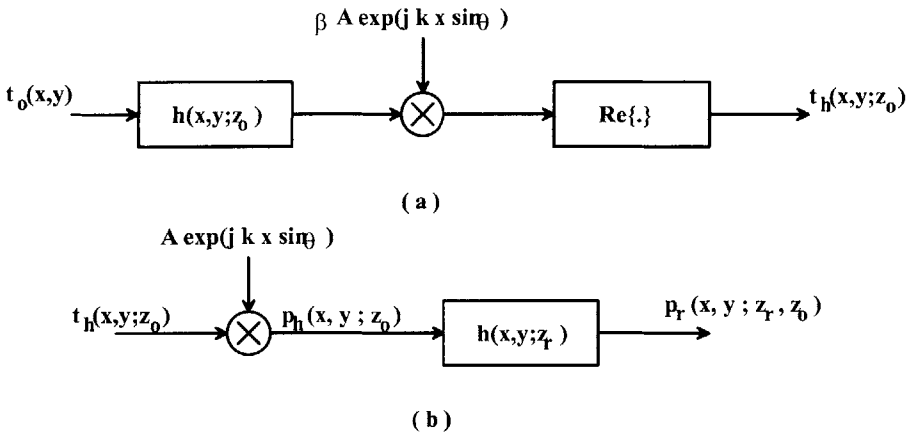

Figure 2-5: Block diagram of general holography recording and reconstruction, after linearization. For in-line holography, $\theta = 0$.
(a) Holography recording process; (b) Holography reconstruction.

The optical field at $z = z_r$, due to the transmitted light p_h at $z = 0$, simply becomes:

$$\begin{aligned} p_r(x_r, y_r; z_r, z_o) &= \{p_h(x, y, z_o) * *h(x, y; z_r)\} \mid_{x=x_r, y=y_r} \\ &= C\{\{\exp(j2kx\sin\theta)[t_o^*(x, y; z_o) * *h^*(x, y; z_o)]\} * *h(x, y; z_r) \\ &\quad +\{[t_o(x, y; z_o) * *h(x, y; z_o)]\} * *h(x, y; z_r)\} \mid_{x=x_r, y=y_r} \end{aligned} \tag{2.23}$$

Where $C = \beta A^2$. Block diagrams corresponding to equation (2.21) and (2.23) are both given in Figure 2-5.

If in addition $z_r = -z_o$, then

$$\begin{aligned} p_r(x_r, y_r; z_r, z_o) &= C\{\{\exp(j2kx\sin\theta)[t_o^*(x, y; z_o) * *h^*(x, y; z_o)]\} * *h(x, y; z_r) \\ &\quad +t_o(x, y; z_o)\} \mid_{x=x_r, y=y_r} \end{aligned} \tag{2.24}$$

Equation (2.24) essentially says that when reconstructed at $z_r = -z_o$, the field shows a focused virtual image and a real image with a different focus which is usually called the twin-image. Both the virtual image and the twin image components are related to the object distribution as shown by the block diagrams in Figure 2-6.

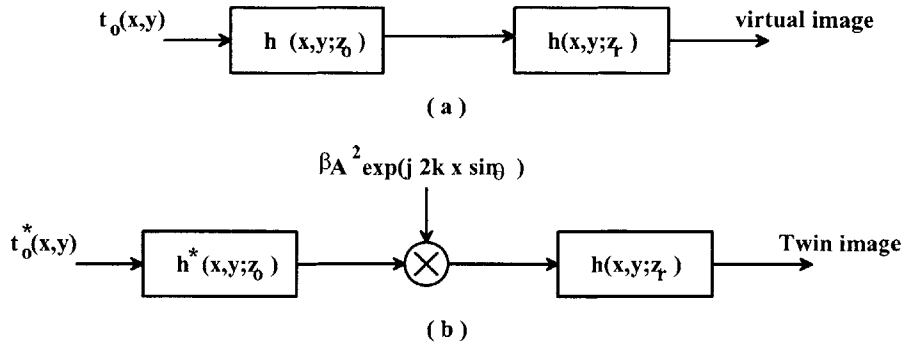

Figure 2-6: Block diagram of the virtual image and the twin-image. For in-line holography, $\theta = 0$.
(a) The virtual image; (b) The twin-image.

Now we consider in-line holography only. Since $\theta = 0$, the twin image overlaps the virtual image in an axial view and manifests itself as an extraneous ghost image. From equation (2.21), if the object distribution is purely absorbing, i.e. $t_o(x, y; z_o)$ is real, then after dropping all constant factors, $t_h(x_h, y_h; z_o)$ can be written as

$$\begin{aligned} t_h(x_h, y_h; z_o) &= \{t_o(x, y) * *Re\{h(x, y; z_o)\}\} \mid_{x=x_h, y=y_h} \\ &= \{t_o(x, y) * *\{\frac{1}{\lambda z} sin[\frac{\pi}{\lambda z}(x^2 + y^2)]\}\} \mid_{x=x_h, y=y_h} \\ T_h(f_x, f_y; z_o) &= T_o(f_x, f_y)H(f_x, f_y) = T_o(f_x, f_y)cos\left[\pi\lambda z_o(f_x^2 + f_y^2)\right] \end{aligned} \tag{2.25}$$

The reconstructed field $p_r(x_r, y_r; z_r, z_o)$ becomes:

$$\begin{aligned} p_r(x_r, y_r; z_r, z_o) &= \{t_h(x, y; z_o) * *h(x, y; z_r)\} \mid_{x=x_r, y=y_r} \\ &= \{t_o^*(x, y; z_o) * *h(x, y; z_r - z_o) \\ &+ \ t_o(x, y; z_o) * *h(x, y; zr + z_o)\} \mid_{x=x_r, y=y_r} \end{aligned} \tag{2.26}$$

2.1.4 Extension To 3D Object Distribution

One might have already found that so far all discussions are based on a disc object model. We have talked about the recording and reconstruction of a hologram associated with an object distribution across one plane, which is located at a fixed distance and parallel to the

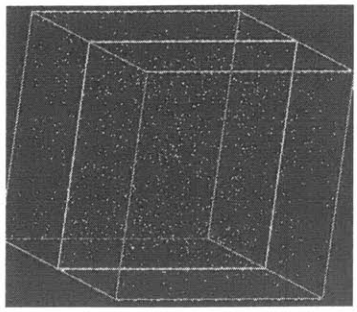

Figure 2-7: A three-dimensional object distribution

hologram. The axial extension of that plane object distribution has not been included in our consideration. However, in real applications, it is often an assembly of small objects, or particles that are involved when a hologram is taken. Each object has physical extension in all directions. Thus a hologram is a record of the wavefield from a three-dimensional object distribution rather than one simple object plane. In such cases, the interpretation of the physical meaning of a hologram and its reconstruction becomes much more complicated than in the disc object cases. Nevertheless, this three-dimensionality is what holography can provide us with.

In this section, we first clarify all the assumptions that we would have to impose upon the object distribution, the wave media and the scattering process, so that for a 3D object distribution, we can still use all the derivations from previous sections based on a disc object model. It will be shown that under these assumptions, a hologram of a 3D object distribution can be approximated by the superposition of holograms contributed by a sequence of disc object distributions. Then an idealistic description of the scattering process throughout the object distribution will be given aiming to explain the information that is contained in a hologram and subsequently expressed by its reconstruction.

First of all, we assume that the object distribution consists of a sparse assembly of small particles contained inside a transparent three dimensional volume. The particles are assumed to be opaque and with some pure absorption. The medium inside the volume is a homogeneous free space other than the presence of particles. The reference light is assumed to be a strong collimated uniform laser beam which has already been polarized.

The scattering occuring in the object distribution is a complicated physical phenomenon. To simplify our discussion, we use assumptions according to the convention [13]. First, we

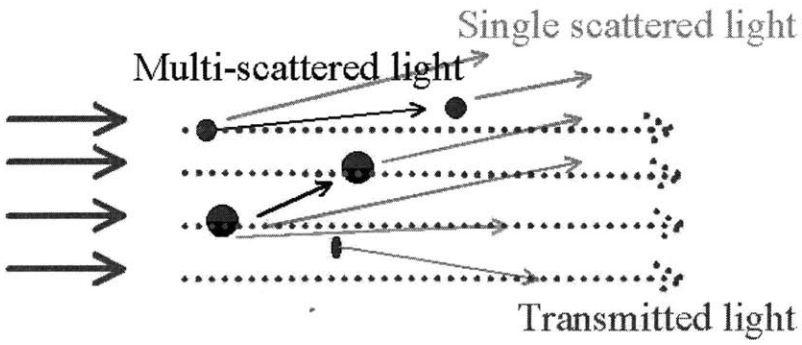

Figure 2-8: Scattering from an object distribution

neglect Muti-scattering. Only the single scattered from each object is considered. Secondly, we neglect the extinction caused to the reference light when it propagates through the medium. Thus all objects are illuminated by the same uniform reference light. Third, all objects are assumed to be have pure absorption, therefore there is no change of polarization and frequency in the scattered light. Finally we assume independent scattering. It means that the phase difference between lights scattered from different objects are not correlated due to the randomness of the distribution. This assumption allows us to sum up all the scattered light in intensity without considering mutual interference.

In summary, all these idealistic assumptions entitle us to consider only two types of wavefield in front of the hologram plane. One is the transmitted reference light which suffers little distortion as we assumed; the other is the total single scattered field contributed by all objects. Therefore, the dominating component of the recorded field becomes the interference between these two waves.

Let's then take a look at the scattering field originated from one object illuminated by the reference light. Figure 2-9 shows that the object body can be divided into three regions associated with different scattering properties: the backward scattering region, the forward scattering region and the geometrical shadow region. Among these three regions, only the information of the forward scattering region reaches the hologram and thus is recorded. We call this region the active region.

Now if we cut the object into many thin slices along the depth direction, as shown in figure 2-10, it is possible to represent the total scattering field by a superposition of contributions from the active region on each slice. At each slice plane, we can use the

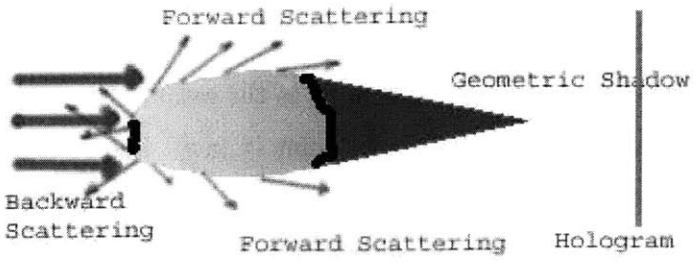

Figure 2-9: Scattering from one object

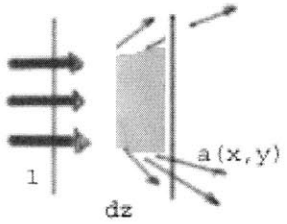

Figure 2-10: The scattering from active regions on one slice of object

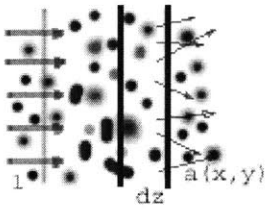

Figure 2-11: The object distribution of one slice of the volume

distribution function representation for the scattering field as in previous sections. We then go back to the large picture of an assemble of particles. We use the same slicing operation which divides the whole volume into many slices. As assumed, the same reference light is assumed for all slices. We then characterize the scattering by objects inside each slice by a distribution function $o(x, y; z)$, which is defined as the ratio of the complex amplitudes at the two sides of the slice in the proximity of point (x, y, z). Figure 2-11 shows a sketch of the slicing operation in a volume.

According to our assumption, the mutual interference between waves from different object slices are independent. The hologram is essentially a linear superposition of their interference with the reference light. Then an hologram of a 3D object space $t_{hv}(x_h, y_h)$ can be represented by integrating $t_h(x, y; z_o)$ along z_o:

$$t_{hv}(x_h, y_h) = \int t_h(x_h, y_h; z_o) dz_o \tag{2.27}$$

Taking t_{hv} for reconstruction, the resulted field can be represented as an integration of $p_r(x_h, y_h; z_r, z_o)$ regarding z_o.

$$p_{rv}(x_h, y_h; z_r) = \int p_r(x_h, y_h; z_r, z_o) dz_o \tag{2.28}$$

When reconstructed at one particular distance z_r, the output consists of one in-focused image of the object plane at z_r and the out of focus image of all other planes, together with all twin-images.

The limits of the validity of this object model depends on the object geometrical shape and scale normalized by the wavelength, and the distance between the object and the hologram as well. A good interpretation of reconstruction images relies on the actual scattering physics.

2.2 Observation and Analysis

From the perspective of computational reconstruction, we are interested in getting high-fidelity recovery of the original object distributions. It will be difficult to define a single deterministic measure for perfect reconstruction. As for our current representation form of the reconstructed field, there are a number of measures for image quality evaluation,

including the transverse and axial resolutions, the degree of distortion and the signal to noise ratio. Note that an image scoring high in terms of any one of these measures is not necessarily the best for human eye's observation. The resolution limit of the reconstructed field is closely tied to the information loss during the whole process. It is especially limited by the phase ambiguity and the effective hologram aperture. For in-line holograms, the noise mainly comes from the twin-image, the intrinsic speckle noise, and from out of focus particles or objects.

In this section, the finite aperture effect and the twin image interference will be discussed. In section 2.2.1, The ideas of the plane wave decomposition and the angular spectrum are briefly reviewed. The finite aperture effect can be described using a convolution expression. In section 2.2.2, in-line holography is described as an analogy to modulation in which the carrier signal is a 2D chirp function. In this sense the twin-image can be represented as a high chirp rate component.

2.2.1 Plane Wave Decomposition, Angular Spectrum Propagation and Finite Aperture Effect

The frequency domain equivalent of equation (2.9) and (2.13) is:

$$O_h(f_x, f_y) = T_o(fx, fy, ; z_o)H(f_x, f_y; z_o) \tag{2.29}$$

Where $O_h(f_x, f_y)$ and $T_o(f_x, f_y; z_o)$ are the Fourier transforms of the diffraction field and the object distribution respectively, or the angular spectrums. According to Fourier Optics theory [10], the angular spectrum of an optical field is its plane wave decomposition, with each spatial frequency pair (f_x, f_y) corresponding to a plane wave propagating in the direction specified by the direction cosine of $(\lambda f_x, \lambda f_y, \sqrt{1-(\lambda f_x)^2-(\lambda f_y)^2})$. Therefore, according to (2.29), the propagation of an optical field simply becomes a dispersive process with a quadratic phase response as specified by equation (2.14), under the paraxial approximation.

Figure 2-12 contains a sketch of an arbitrary angular spectrum in one direction as well as the components of the spatial spectra of its in-line and off-axis holograms. If the spectrum of the halo is small in comparison to the spectrum of the objects being encoded, the spatial sampling frequency of the in-line hologram must be twice the highest frequency

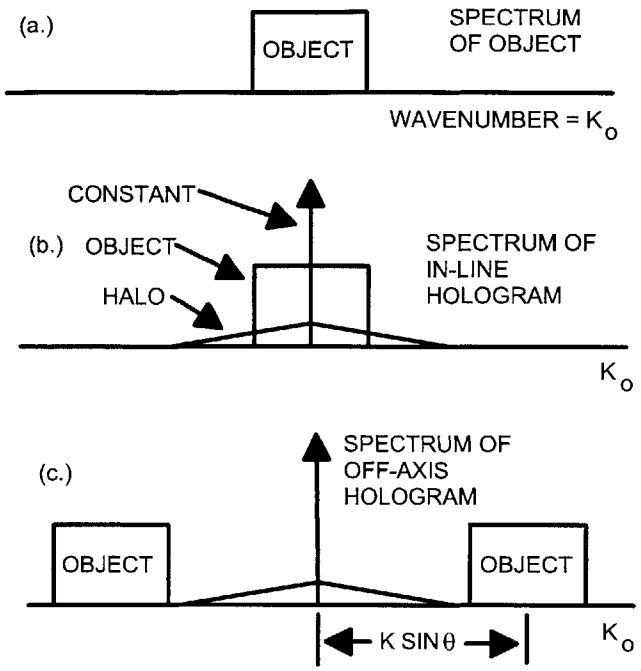

Figure 2-12: Spectra in one dimension for an image and its in-line and off-axis holograms a. Image, b. In-line hologram, c. off-axis hologram

in the image to satisfy the sampling theorem. It seems that only a modest advantage from a higher sampling frequency since the halo represents "noise" whether or not it is aliased. In practice, the relative intensity of the halo is minimized by making the reference beam intensity large in comparison to the object beam intensity, when possible. For an off-axis hologram, if the off-axis angle is large enough to separate the desired phase encoding from the halo the required sampling frequency is four times larger than the required in-line frequency, and if the angle allows halo noise, but just separates the twin images, the required frequency is two times the needed in-line frequency. These increases apply to only one of the two directions on the hologram since the off-axis modulation is in only one direction.

Though $H(f_x, f_y; z_o)$ is an all-pass filter which is neither band limited nor space limited, the actual recordable field is limited in spatial extension by the hologram size and reference beam span. Therefore information loss happens when an hologram is recorded. For simplicity, first we define an effective hologram aperture as the minimum of the reference beam span and the physical format of the recording medium. This effective hologram aperture stands for the maximum spatial extension of the recordable holographic information. It can

be included into consideration by simply imposing a window on the hologram. That is

$$t_h(x_h, y_h; z_o) = w(x_h, y_h)\{t_o(x, y) * *Re\{h(x, y; z_o)\}\} \mid_{x=x_h, y=y_h} \tag{2.30}$$

where $w(x_h, y_h)$ is the window function. It can be circular symmetric:

$$w(x_h, y_h) = \begin{cases} 1 & \sqrt{x_h^2 + y_h^2} \leq R \\ 0 & \text{otherwise} \end{cases}$$

Or in a rectangular form:

$$w(x_h, y_h) = \begin{cases} 1 & \text{if } x_h \leq R \text{ and } y_h \leq R \\ 0 & \text{otherwise} \end{cases}$$

Now let's check the field reconstructed from the windowed hologram. After including the effective finite aperture window, equation (2.26) becomes

$$\begin{aligned} p_r(x_r, y_r; z_r, z_o) &= \{\{w(x, y)[t_o(x, y) * *Re\{h(x, y; z_o)\}]\} * *h(x, y; z_r)\} \mid_{x=x_r, y=y_r} \\ &= \frac{1}{2}\{\{w(x, y)[t_o(x, y) * *h(x, y; z_o)]\} * *h(x, y; z_r)\} \mid_{x=x_r, y=y_r} \\ &+ \frac{1}{2}\{\{w(x, y)[t_o(x, y) * *h^*(x, y; z_o)]\} * *h(x, y; z_r)\} \mid_{x=x_r, y=y_r} \\ &= u_1 + u_2 \end{aligned} \tag{2.31}$$

Where u_1 and u_2 denote the virtual and real image component respectively.

We are mainly concerned about the effect of this finite aperture on the focused virtual image. Since the twin-image component u_2 is an interference term anyway, we only discuss the effect on u_1 here. Expressed in an integral form, u_1 becomes:

$$\begin{aligned} u_1 &= \frac{1}{2}\iint_{-\infty}^{\infty} w(x_h, y_h) \iint_{-\infty}^{\infty} t_o(x_o, y_o) h(x_h - x_o, y_h - y_o; z_o) dx_o dy_o \; h(x_r - x_h, y_r - y_h; z_r) dx_h dy_h \\ &= \frac{1}{2}\iint_{-\infty}^{\infty} t_o(x_o, y_o) \iint_{-\infty}^{\infty} w(x_h, y_h) h(x_h - x_o, y_h - y_o; z_o) h(x_r - x_h, y_r - y_h; z_r) dx_h dy_h dx_o dy_o \end{aligned} \tag{2.32}$$

Let $z_r = -z_o$ and substitute $h_f(x, y; z) = \frac{1}{j\lambda z} \exp j\frac{\pi}{\lambda z}(x^2 + y^2)$ into above, u_1 becomes:

$$
\begin{aligned}
u_1 &= \frac{1}{2\lambda^2 z_o^2}\iint_{-\infty}^{\infty} t_o(x_o,y_o)\exp\{j\frac{\pi}{\lambda z_o}(x_o^2-x_r^2+y_o^2-y_r^2)\} \\
&\quad \{\iint_{-\infty}^{\infty} w(x_h,y_h)\exp\{j\frac{2\pi}{\lambda z_o}[x_h(x_r-x_o)+y_h(y_r-y_o)]\}dx_h dy_h\}dx_o dy_o \\
&= \iint_{-\infty}^{\infty} t_o(x_o,y_o)\exp\{j\frac{\pi}{\lambda z_o}(x_o^2-x_r^2+y_o^2-y_r^2)\}W(\frac{2\pi}{\lambda z_o}(x_r-x_o),\frac{2\pi}{\lambda z_o}(y_r-y_o))dx_o dy_o
\end{aligned}
\tag{2.33}
$$

Where $W(\frac{2\pi}{\lambda z}(x_r-x_o),\frac{2\pi}{\lambda z}(y_r-y_o))$ is the 2D Fourier transform of the window function $w(x_h,y_h)$. If the window size is infinitely large, i.e. $w(x_h,y_h)=1$, then $W(\frac{2\pi}{\lambda z}(x_r-x_o),\frac{2\pi}{\lambda z}(y_r-y_o))$ becomes $\delta(x_r-x_o)\delta(y_r-y_o)$, therefore,

$$
\begin{aligned}
u_1 &= \frac{1}{2}\iint_{-\infty}^{\infty} t_o(x_o,y_o)\delta(x_r-x_o)\delta(y_r-y_o)dx_o dy_o \\
&= \frac{1}{2}t_o(x_r,y_r)
\end{aligned}
\tag{2.34}
$$

which means the virtual image is a perfect reconstruction. However when $w(x_h,y_h)$ has finite support as in all real applications, the reconstructed virtual image is not a perfect reconstruction of the original object distribution. Rather it is an phase modulated object distribution convolved with the Fourier transform of the windows function. When the window size is large relative to $\sqrt{\lambda z}$, $W(\frac{2\pi}{\lambda z}(x_r-x_o),\frac{2\pi}{\lambda z}(y_r-y_o))$ is very close to $\delta(x_r-x_o)\delta(y_r-y_o)$, thus the finite aperture effect is negligible. With the finite aperture effect included, block diagrams of virtual image and twin-image should be modified into Figure 2-13.

2.2.2 Chirp Carrier Modulation

In this section, we will describe the twin-image effect of in-line holography using an analogy to the well-known amplitude modulation(AM) process. We will show that due to the natural of the twin-image component, it is very difficult to separate the virtual image and the twin-image using traditional filtering techniques, if not impossible. Recall that in communications an AM signal $s(t)$ is generated by multiplying a modulating signal $a(t)$ with a carrier signal

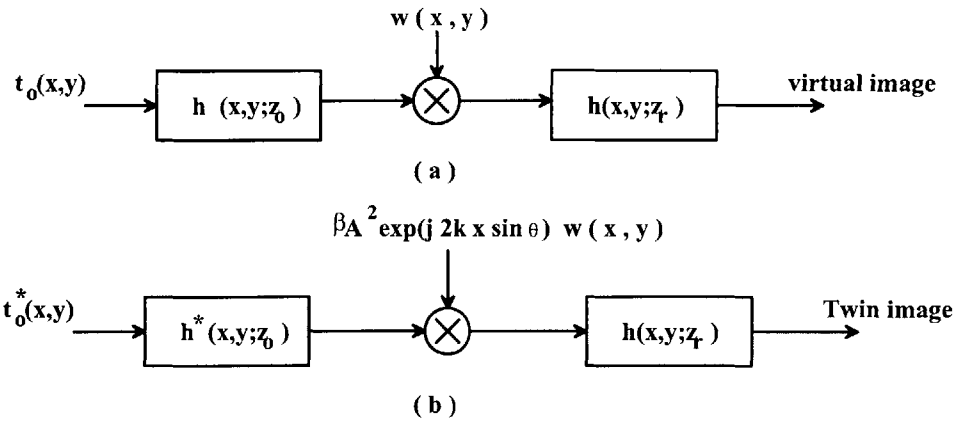

Figure 2-13: Block diagram of the virtual image and the twin-image, with finite aperture effect. For in-line holography, $\theta = 0$.
(a) The virtual image; (b) The twin-image

$c(t) = cos(w_c t)$, w_c is the carrier frequency. That is

$$s(t) = a(t)cos(w_c t) \tag{2.35}$$

At the receiver end, the received signal is $r(t) = s(t) + n(t)$, where $n(t)$ is the additive noise. An estimation of $a(t)$ can be extracted from $r(t)$ by demodulation, that is, first multiplying $r(t)$ by the same carrier signal $c(t)$ and then doing low-pass filtering, as shown below:

$$\begin{aligned} \hat{a}(t) &= \{r(t)cos(w_c t)\} * h(t) \\ &= \{a(t) * cos^2(w_c t) + n(t) * cos(w_c t)\} * h(t) \\ &= \{\frac{1}{2}a(t) + \frac{1}{2}a(t)cos(2w_c t) + n(t)cos(w_c t)\} * h(t) \end{aligned} \tag{2.36}$$

Where $h(t)$ is the impulse response of the low pass filter, so that the baseband components including $a(t)$ can be separated from the high-frequency component $a(t)cos(2w_c t)$.

Recall the expressions of the hologram $T_h(f_x, f_y; z_o)$ and its in-focus reconstruction $p_r(f_x, f_y; -z_o, z_o)$ given in section 2.1.3, their forms are analogous to those of $s(t)$ and $\hat{a}(t)$. This suggests an analogy between in-line holography and AM modulation. In equation (2.25), the spatial spectrum of an hologram is the product of the angular spectrum of the object and $cos\left[\pi\lambda z_o(f_x^2 + f_y^2)\right]$, the angular spectrum of the kernel. The kernel can be thought of as the carrier signal and the object spectrum as the modulating signal. Therefore

in this sense the recording of a hologram becomes a modulation process in angular frequency domain. Since the carrier signal here is a chirp function, we call it chirp carrier modulation. The reconstructed field, as formulated in equation (2.26), is equivalent to the demodulated output. Therefore, the two steps process can be described in frequency domain as:

1. Modulation.

$$T_h(f_x, f_y; z_o) = 2T_o(f_x, f_y)cos\left[\pi\lambda z_o(f_x^2 + f_y^2)\right] \tag{2.37}$$

2. Demodulation.

$$\begin{aligned} p_r(f_x, f_y; z_r = -z_o) &= T_h(f_x, f_y; z_o)\exp\{j\left[\pi\lambda z_r(f_x^2 + f_y^2)\right]\} \\ &= 2T_o(f_x, f_y)cos\left[\pi\lambda z_o(f_x^2 + f_y^2)\right]\exp\{-j\left[\pi\lambda z_o(f_x^2 + f_y^2)\right]\} \\ &= T_o(f_x, f_y) + T_o(f_x, f_y)\exp\{-j\left[2\pi\lambda z_o(f_x^2 + f_y^2)\right]\} \end{aligned} \tag{2.38}$$

From equation (2.38), we can see that the twin-image is a chirp signal whose chirp rate is twice of that of the carrier signal. This is similar to the high-frequency term we got in AM demodulation. Though the whole process is analogous to AM modulation and demodulation, holography is different from AM in several aspects:

First, unlike in AM where single tone carriers are often used, in in-line holography the equivalent carrier signal is a 2D chirp signal with a quadratic phase structure.

The second difference is a direct consequence of the first one. In AM, the demodulated original signal is separable from other components in frequency domain and thus can be extracted simply by low-pass filtering. Unfortunately this is not the case for in-line holography. The result in equation (2.38) indicates that the two components are separable neither in spatial domain nor in frequency domain. The two components are different only in their chirp rate(see definition in section 2.1.2): $C_r = 0$ for the virtual image and $C_r = 2/(\lambda z_o)$ for the real twin image. If there is a way to transform p_r into the chirp rate domain and if the transform is invertible, then the two terms can be separated by doing filtering in the C_r domain.

To visualize this analogy, Figure 2-14 and Figure 2-15 have illustrated the signals, their spectrums and spectrograms in an AM process and an in-line holography process respectively. Figure 2-14a, b, c show the waveforms, spectrums and spectrograms of the modulat-

ing signal $a(t)$ respectively. $a(t)$ is a sinc function truncated in time. Its spectrum is a box function convolved with a sinc function in frequency. Therefore, all its energy is concentrated along the base-band region in the spectrogram. Figure 2-14d, e, f have plotted the waveforms, spectrums and spectrograms of the amplitude modulated signal. The carrier frequency has been chosen to be 1/8 of the sampling frequency. In the spectrogram, the energy can be found to be maximized along $f = 1/8 f_s$. The demodulated signal, without considering any additive noise, is shown in figure 2-14g, together with its spectrum and spectrogram in figure 2-14h and i respectively. Compositions of the demodulated signal, including the baseband signal and the high-frequency component, can be easily observed from these images, and are well separated in frequency domain.

An equivalent set of results has been computed for in-line holography and shown in Figure 2-15. Note that it is only a one-dimensional simplified version. A same modulating signal has been used as in the AM case, except that now it is a sinc function in spatial frequency. The carrier signal has a chirp rate $C_r = 1.58 * 10^7\ m^{-2}$ or $6.1035 * 10^{-5} f_s\ m^{-1}$. The modulated signal is shown in Figure 2-15d, and its spectrum and spectrogram are plotted in Figure 2-15e, f respectively. The spectrum is still found to be centered on the baseband. The energy distribution in the spectrogram as shown lies along two symmetric shear lines, each having a slope of C_r and $-C_r$ respectively. The waveform, spectrum and spectrogram of the demodulated signal are shown in Figure 2-15g, h and i respectively. In the spectrogram image, the virtual image component is the straight line lying along $f = 0$. The twin image component corresponds to the shear line with slope of $-2C_r$. However, when projected onto either the spatial axis or the frequency axis, the two components overlap each other. It may be possible to separate them by spatial-frequency joint process methods since we can clearly tell them one from the other in the spectrogram.

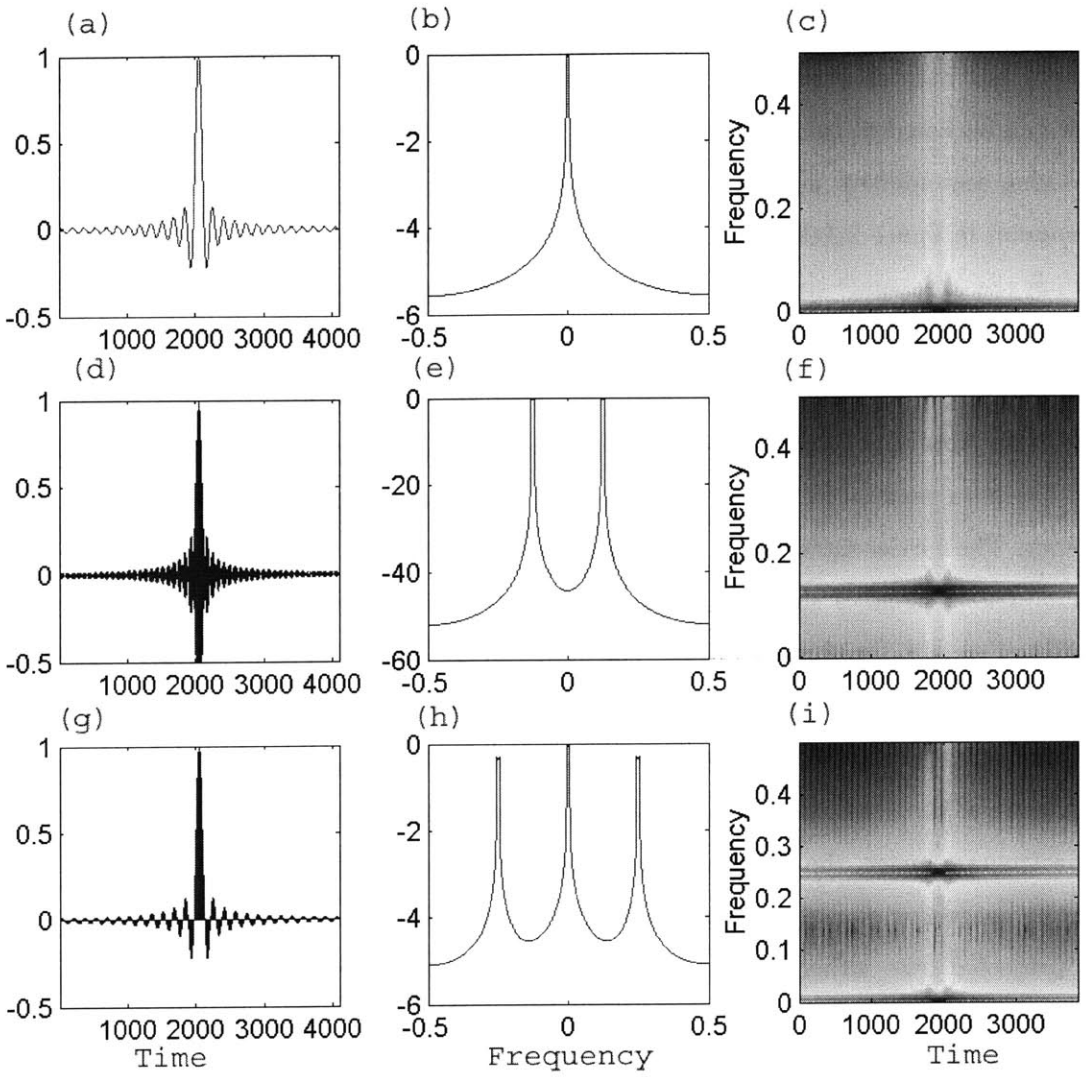

Figure 2-14: AM signals, spectrums and spectrograms
(a) $a(t) = sinc(50tf_s)rect(2048/f_s)$. (b) The normalized spectrum of a(t).
(c) The Spectrogram of $a(t)$(Window length is $256/f_s$, step is $10/f_s$)
(d) The modulated signal $s(t) = a(t)cos(2\pi f_s/8t)$. (e) Spectrum of $s(t)$.
(f) The Spectrogram of $s(t)$(the same window length and step as in (c))
(g) The demodulated signal $r(t) = s(t)cos(2\pi f_s/8t)$. (h) Amplitude spectrum of $r(t)$.
(i) The Spectrogram of $r(t)$(the same window length and step as in (c))

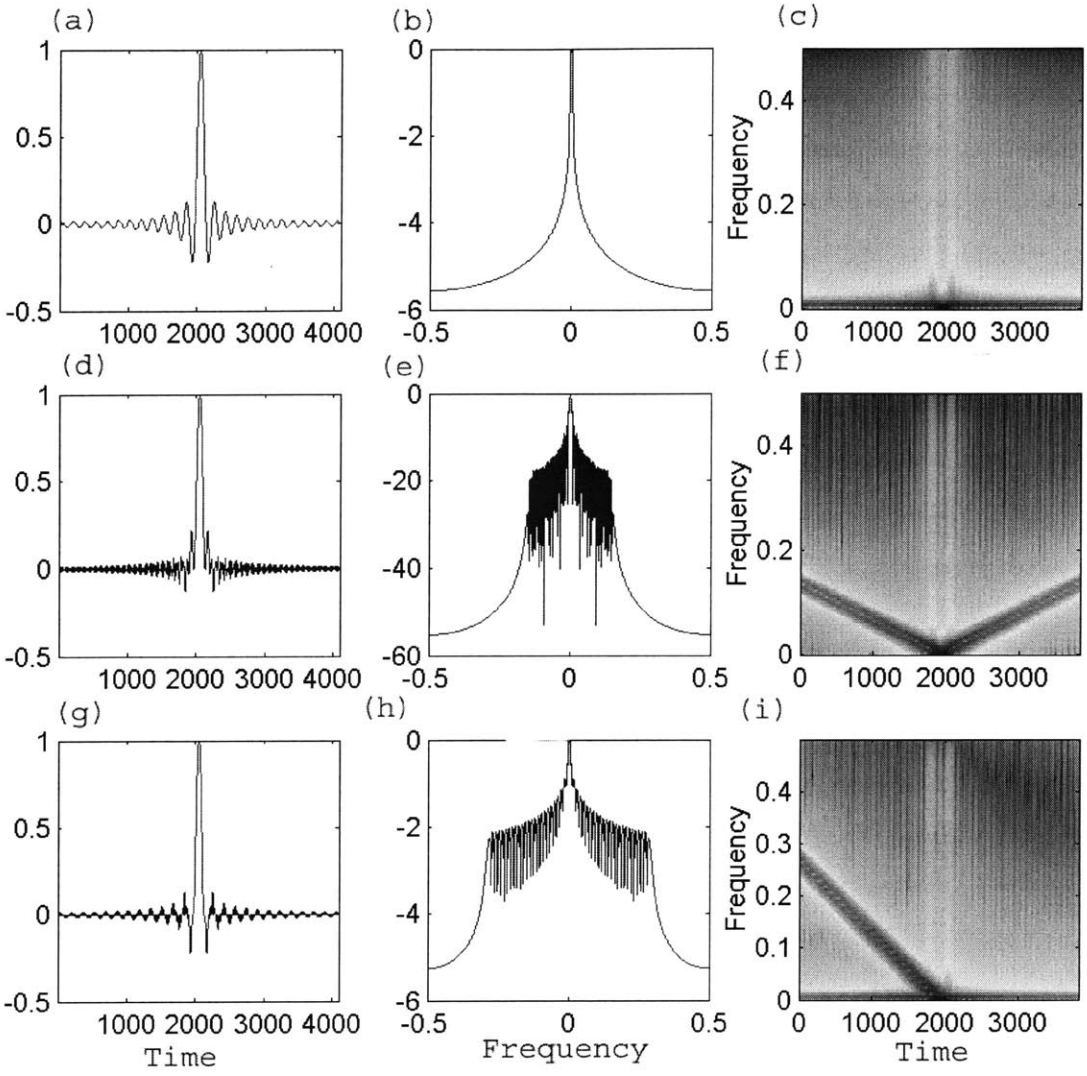

Figure 2-15: 1D holographic signals, spectrums and spectrograms
(a) $a(f) = sinc(50fL)rect(2048/L)$. (b) The normalized spectrum of $a(f)$.
(c) The Spectrogram of $a(f)$(Window length is $256/L$, step is $10/L$)
(d) The modulated signal $s(f) = a(f)cos(\pi\lambda z_0 f^2)$, $\lambda = 633nm$, $z_0 = 100mm$.
(e) The spectrum of $s(f)$. (f) The Spectrogram of $s(f)$(the same parameters as (c))
(g) The demodulated signal $r(f) = s(f)exp(\pi\lambda z_r f^2)$, $z_r = -z_0$.
(h) Amplitude spectrum of $r(f)$.
(i) The Spectrogram of $r(f)$(the same parameters as in (c))

Chapter 3

Computational Reconstruction of In-Line Holograms

Starting from this chapter, our attention will be mainly focused on the computational reconstruction of in-line holograms. Computational reconstruction of an object distribution from its hologram is essentially an inverse problem. The formulation of the forward problem has been given by (2.25) in section 2.1.3. Unfortunately, a direct inverse does not exist since the kernel $Re\{h(x,y;z_o)\} = cos\left[\pi\lambda z_o(f_x^2+f_y^2)\right]$ has infinitely many zeros. The optical reconstructed field is not an exact inverse, rather it is a pseudo-inverse to the first order, which defectively has the twin-image artifacts. Previous work includes a higher order pseudo inverse developed by Onural [19] which as stated can effectively reduce the twin-image effect at small order numbers. As terms of order number higher than 5 being included, extra noise will be introduced due to the zeros of the forward kernel. Also aliasing artifacts increase dramatically as the order number rises. Also the problem has been treated from another different path. Note that the twin-image is equivalently an ambiguity due to the loss of total interaction field phase (or the loss of the phase sign of the object waves)when the hologram is recorded, thus the reconstruction cannot discriminate the conjugate wavefront curvature from the original wave curvature. In this sense the hologram is merely an magnitude mapping of the total interaction field. If the phase structure of the total field could be determined, perfect reconstruction would be achieved without any ambiguity. This idea leads to the so called phase retrieval method. The basic idea is to retrieve the wavefront phase information from several magnitude records of the wavefield. It has been widely ap-

plied in many different applications including holography [9][14]. In 1971, R.W.Gerchberg and W.O.Saxton proposed an iterative phase-retrieval algorithm [9] which determines the phase from the intensity distributions of the object plane and a diffraction plane. The algorithm starts with a random phase structure as the initial value. The final solution is unique as they proved, due to the Fourier transform constraint between these two intensity distributions. However, the application of Gerchberg-Saxton algorithm is limited by its requirement of the known object plane magnitude which is rather an unknown to be determined in our case. This requirement was later waived by G. Liu [14]. He started with a conventional reconstruction. Then the reconstructed field was feedback to provide an estimation of the magnitude and phase structure of the object distribution. This method has a better convergence rate than Gerchberg-Saxton algorithm However, it still converges slowly and involves large amount of computations since 2D FFT and 2D inverse FFT are both required at each iteration. This is formidable for large holograms.

In this chapter, we first present computational methods of implementing the optical reconstruction process. The behavior of this implementation is then described in terms of several focus measures. Based on a simple circular object distribution function, focus measure curves have been computed and shown. They are associated with the MSE error, the energy concentration and the maximum gradient of the reconstructed images. Then modified kernel functions are proposed to improve the reconstructed image quality, namely by spatial windowing and filtering. They both equivalently correspond to high-pass filtering. By calculating reconstructions from a selective band of angular spectrum of a hologram, the influence of the directly transmitted light can be reduced which dominates the low-frequency region. The spatial windowing method can be related to the so called "In-Line Recording and Off-Axis Viewing (IROV)" technique which was proposed by H. Meng [17].

Specifically, our computational procedure can be visualized as shown in figure 3-1. First, the system takes a digitized optical hologram or a digital hologram image as the input. Then reconstruction of the associated object distribution proceeds plane by plane sequentially inside the object space, with each plane parallel to the hologram and perpendicular to the optical axis. The computational reconstruction of each plane is parameterized by its distance to the hologram. The distance can be arbitrarily chosen inside the object space and the generated image is an estimate of the object field distribution across the plane at that distance. Therefore the reconstructed 3D object space is essentially represented by its

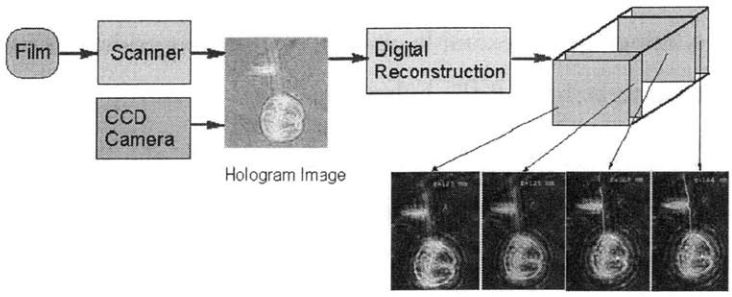

Figure 3-1: The Diagram of computational reconstruction

cross sections sampled along the axial direction, a way very similar to what CT-scanning does. Once these 2D images have been computed, there are a variety of techniques that can be applied to assimilate them into a 3D reconstruction. This post processing step is outside the scope of holographic reconstruction and won't be discussed in this thesis.

The content of this chapter is then organized as the following: First, an optimal sampling scheme of holograms is proposed in Section 3.1. In section 3.2, Computational implementation of optical reconstruction is discussed, which is directly based on the formulation in section 2.1.3. Issues involved in discretization ad truncating are discussed, including discrete kernel construction. In section 3.2.3, the behavior of this implementation is described as a matched filtering and a number of focus measures have been proposed. Then in section 3.3, we revise the Fresnel kernel both in the spatial domain and the frequency domain, aimed to improve the reconstruction image quality. These modifications are equivalent to high-pass filtering. Examples of reconstructed images with and without filtering are shown.

3.1 Optimal Sampling of Holograms

Holographic recording media, regardless of their forms, are essentially sampling devices associated with certain physical resolutions. To be accessible to digital computers, however, optical holograms have to be discretized and quantized into digital hologram images, before they can be computationally reconstructed. The sampling of a hologram is more than just a routine scan of a transparency or pictures. The sampling scheme is crucial to reconstruction quality and efficiency. Efficiency stands for the computing size and speed; quality refers to image resolution and intelligibility. When reconstructed optically, a hologram is equivalent

to a diffraction grating and the diffraction field converges to the reconstructed images. The effect of sampling a hologram, for example, using a rectangular grids, is equivalent to putting a lattice grating mask right behind the hologram.

Similar to the Nyquist sampling theorem for time signals, the Whittaker-Shannon sampling theorem applies for 2D hologram sampling [10]. It states that from a rectangular lattice of samples of a 2D function $h(x, y)$:

$$h_s(x, y) = comb(\frac{x}{X})comb(\frac{y}{Y})h(x, y) \tag{3.1}$$

it is possible to successfully recover the original data $h(x, y)$ if the samples are taken close enough such that their separations are smaller than two maximum spacings:

$$X \leq \frac{1}{2B_x}, \quad Y \leq \frac{1}{2B_y} \tag{3.2}$$

where B_x and B_y are the bandwidth in f_x and f_y directions, X and Y are sampling steps along x and y directions respectively, and

$$comb(x) = \sum_{n=-\infty}^{\infty} \delta(x - n) \tag{3.3}$$

Therefore, in order to determine X and Y, it is necessary to know B_x and B_y, the bandwidth of the hologram in f_x and f_y directions. If the hologram is assumed infinitely large, then its bandwidth would be solely determined by the bandwidth of the object wave. However, in real applications, both the recording media and reference laser beam span are of finite size. Therefore, a hologram often has a more restrictive bandwidth.

For simplicity, we first derive the bandwidth in one dimensional case. Generalization to the 2D cases is straightforward. As illustrated in Figure 3-2, in the y direction the object has length L_{oy} and the hologram has length L_{hy}. Both are centered on the optical axis, i.e. z axis. We assume that the incident reference light is collimated and parallel to the optical axis, with a beam span larger than the hologram size. Since the interference between the reference light and the object field dominates the total interaction field in front of the hologram plane, we can just consider the bandwidth of that interference field, which is then determined by the largest directional angle difference between two interference sources. Since the reference light has a fixed direction parallel to the optical axis, the bandwidth of

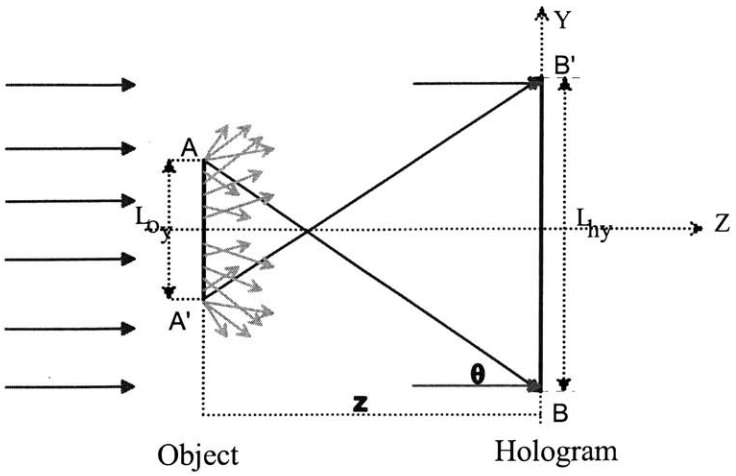

Figure 3-2: The bandwidth of a hologram of finite format;

the interference field is equal to the bandwidth of the object field. The bandwidth of the object, or its angular spectrum, consists of numerous plane waves propagating in a diverse directions. However, it is only those falling on the hologram plane that will be recorded as they interfere with the reference light. The largest spatial frequency corresponds to the plane wave emitted from point A (or A')on the object arriving on point B (or B') on the hologram. It can be related to L_{oy}, L_{hy} and z by:

$$f_{ymax} = \frac{sin\theta}{\lambda} = \frac{\frac{L_{oy}+L_{hy}}{2}}{\lambda\sqrt{(\frac{L_{oy}+L_{hy}}{2})^2 + z^2}} \tag{3.4}$$

If we assume further that $L_{oy} + L_{hy} \ll z$, then $\sqrt{(\frac{L_{oy}+L_{hy}}{2})^2 + z^2} \approx z$, and

$$f_{ymax} = \frac{L_{oy} + L_{hy}}{2\lambda z} \tag{3.5}$$

Similarly, if the object length and hologram length in x direction are L_{ox} and L_{hx} respectively, then

$$f_{xmax} = \frac{L_{ox} + L_{hx}}{2\lambda z} \tag{3.6}$$

In summary, the hologram bandwidth in f_x and f_y directions are:

$$\begin{aligned} 2B_y &= 2f_{ymax} = \frac{L_{oy} + L_{hy}}{\lambda z} \\ 2B_x &= 2f_{xmax} = \frac{L_{ox} + L_{hx}}{\lambda z} \end{aligned} \tag{3.7}$$

The same results can be inferred by thinking of the hologram as a set of fringe pattern resulting from the interference between two plane waves. Sampling the hologram is then equivalent to the sampling of the variation of the interference patterns. The bandwidth of these fringe patterns is determined by the angle between the wavenumber vectors of the two interference sources.

The above relationships coincide with Goodman's results in [10], where they are derived from the Fresnel transform representation. The Fresnel transform can be written in a form of modified Fourier transform:

$$\begin{aligned} t_h(x_h, y_h; z_o) &= \exp\{j\frac{\pi}{\lambda z}(x_h^2 + y_h^2)\} \\ &\iint t_o(x_o, y_o) \exp\{j\frac{\pi}{\lambda z}(x_o^2 + y_o^2)\} \exp\{-j\frac{2\pi}{\lambda z}(x_h x_o + y_h y_o)\} dx_o dy_o \end{aligned} \tag{3.8}$$

The bandwidth of t_h is equal to the bandwidth of the quadratic complex exponential $\exp^{j\frac{\pi}{\lambda z}(x_h^2+y_h^2)}$ plus the bandwidth of the integration. First, the bandwidth of the complex exponential can be approximated by its local frequency as given in section 2.1.2, which is $\frac{L_{hy}}{\lambda z}$. On the other hand, the integration is nothing but a Fourier transform of the object multiplied by a quadratic complex exponential. Hence its bandwidth can be approximated by $\frac{L_{oy}}{\lambda z}$ since the complex exponential does not change the spatial extension of t_o [10]. Therefore, the total bandwidth becomes $\frac{L_{oy}+L_{hy}}{\lambda z}$.

Combine (3.7) with (3.2), the optimal sampling steps in x and y directions can be derived as:

$$X = \frac{\lambda z}{L_{ox} + L_{hx}}, \quad Y = \frac{\lambda z}{L_{oy} + L_{hy}} \tag{3.9}$$

And the total number of samples in each dimension becomes

$$NX = \frac{L_{hx}(L_{ox} + L_{hx})}{\lambda z}, \quad NY = \frac{L_{hy}(L_{oy} + L_{hy})}{\lambda z} \tag{3.10}$$

Finally, the sampled hologram image is represented by

$$t_h[m,n] = t_h(mX, nY), m = -NX/2, ..., NX/2 - 1 \text{ and } n = -NY/2, ..., NY/2 - 1 \quad (3.11)$$

The suggested sampling scheme is optimal in the sense that it satisfies the Whittaker-Shannon sampling theorem with the minimum computation cost. Note that it is z dependent, which means that to reconstruct objects at different distances from one hologram, the same hologram needs to be re sampled for the reconstruction of objects at each distance. This is quite cumbersome practically and usually not used. In practical implementations, a hologram is only sampled once at the highest possible resolution. The problem is that, the values of NX and NY as given in (3.10) are often so significantly large that the corresponding resolution requirement is impractical to achieve. For example, assume that the object and hologram have the same size $0.025 * 0.025\ m^2$, and they are separated by $0.1\ m$, also the wavelength is $\lambda = 633e - 9\ m$, then $NX = NY = 19747$ and $X = Y = 1.27\mu m$. Problems raised here are not only that the amount of data involved is too substantial, but also that the resolution is too high to be attainable by most commercial scanners. The situation becomes worse as the distance reduces. Fortunately the results given in (3.9) and (3.10) are concluded from the worst case consideration. When the object distribution is sparse and individual particles are small, for reasonable image quality the effective object size does not have to be the whole aperture. Hence compromises can be made to reduce the total number of sampling, while the resulting image quality can still be satisfying in most cases.

A number of other papers have also been published to address this problem [11] [24] [20]. But none of these took into consideration the finite aperture effects of both the object and hologram planes.

3.2 Computational Implementation of Optical Reconstruction

3.2.1 Kernel Construction

The spatial kernel, according to section 2.1.3, is $h_e(x, y; z)$ without the paraxial approximation, or is the Fresnel kernel $h_f(x, y; z)$ with paraxial approximation. They are both

parameterized by λz. The reconstructed optical field at a distance z is the output of convolving the hologram with one of these kernels. Often the 2D convolution in spatial domain is computed via multiplications in frequency domain by using FFT. Therefore the Fourier transform of each spatial kernel needs to be calculated. For the Fresnel kernel $h_f(x, y; z)$, its transform $H_f(f_x, f_y; z)$ is analytically known, thus it can be constructed directly in frequency domain. The Fourier transform of the exact kernel $h_e(x, y; z)$, however, can not be found analytically. In this case we have to calculate $h_e(x, y; z)$ first, and then compute the Fourier transform $H_e(f_x, f_y; z)$ numerically by computing the 2D FFT of $h_e(x, y; z)$ for each value of z.

The kernels $h_e(x, y; z)$, $h_f(x, y; z)$ and $H_f(f_x, f_y; z)$ are re-written as follows:

$$h_e(x, y; z) = \frac{1}{j\lambda} \frac{exp(jk\sqrt{x^2 + y^2 + z^2})}{\sqrt{x^2 + y^2 + z^2}} \tag{3.12}$$

$$h_f(x, y; z) = \frac{1}{j\lambda z} exp[\frac{jk}{2z}(x^2 + y^2)] = \frac{1}{j\lambda z} exp[\frac{j\pi}{\lambda z}(x^2 + y^2)] \tag{3.13}$$

$$H_f(f_x, f_y; z) = \mathcal{F}\{h_f(x, y; z)\} = \exp\left[-j\pi\lambda z(f_x^2 + f_y^2)\right] \tag{3.14}$$

For computational implementations they have to be discretized and truncated. Both the sample rate and appropriate kernel size needs to be determined.

Spatial Domain Kernel

We first discuss the construction of the discrete spatial kernels: $h_e(x, y; z)$ and $h_f(x, y; z)$. First the same sample rate has to be used as that of the hologram. The second step is to determine the kernel size. The answer is not obvious at the first look, since the so called kernel function is rather a mathematical abstraction derived from the modeling of optical reconstruction. There is no obvious way to figure out what the right size for the kernel should be even the hologram size is known. It is true that the hologram size cannot solely decide the kernel size which, as we will show, is actually distance dependent. Recall that the Fresnel kernel given in section 2.1.2, its local frequency in x or y direction increases linearly with the associated dimension. The exact kernel can also be approximated by this. Hence if the kernel size increases without limit, the highest local frequency will pass over half the sample frequency and then aliasing will happen. Therefore to avoid aliasing the appropriate kernel size must be selected such that the largest local frequency is equal

to half the sample frequency. Consequently, larger kernel size is required to get higher reconstruction resolution. This implies higher sample frequency for both the kernel and hologram. Let $L_{kx} * L_{ky}$ be the kernel size, f_{xs} and f_{ys} the sample frequencies in each dimension, and X and Y the sample spacings in each dimension, then we have the following relations:

$$\begin{aligned} f_{xs} &= \frac{1}{X}, \quad f_{ys} = \frac{1}{Y} \\ f_x(\frac{L_{kx}}{2}) &= \frac{L_{kx}}{2\lambda z}, \quad f_y(\frac{L_{ky}}{2}) = \frac{L_{ky}}{2\lambda z} \\ L_{kx} &= \frac{\lambda z}{X}, \quad L_{ky} = \frac{\lambda z}{Y} \\ N_{kx} &= \frac{L_{kx}}{X} = \frac{\lambda z}{X^2}, \quad N_{ky} = \frac{L_{ky}}{Y} = \frac{\lambda z}{Y^2} \end{aligned} \tag{3.15}$$

If the hologram is sampled at optimal sampling rate given in (3.9), the kernel size is equal to the geometric average of the object size and the hologram size, i.e.

$$\begin{aligned} L_{kx} &= \frac{L_{ox} + L_{hx}}{2} \\ L_{ky} &= \frac{L_{oy} + L_{hy}}{2} \end{aligned} \tag{3.16}$$

Note that this is not true if the sampling of hologram is not done according to (3.9). And if the hologram is sampled under the optimal rate according to (3.9), the spatial kernel size can still be inferred from (3.15), which decreases as the sample spacing increase.

Based on above, the discrete version of each spatial kernel is ready to be written down as

$$\begin{aligned} h_f[m,n;z] &= h_f(mX,nY;z) = \frac{1}{j\lambda z}\exp\{j\frac{\pi}{\lambda z}[(mX)^2+(nY)^2]\} \\ h_e[m,n;z] &= h_e(mX,nY;z) = \frac{z}{j\lambda}\frac{\exp(j\frac{2\pi}{\lambda}\sqrt{(mX)^2+(nY)^2+z^2})}{(mX)^2+(nY)^2+z^2} \end{aligned} \tag{3.17}$$

where $m = -N_{kx}/2, ..., N_{kx}/2-1$ and $n = -N_{ky}/2, ..., N_{ky}/2-1$.

Frequency Domain Kernel

By the frequency domain kernel, we refer to the Fourier transform of the Fresnel kernel, $H_f(f_x, f_y; z)$, which is analytically available. Thus it is possible to construct it directly in

the frequency domain without doing FFT's. $H_f(f-x, f_y; z)$ also needs to be sampled and truncated for computational implementations, but in this case in the frequency domain. As in the spatial case, we need to determine the sample spacing and the appropriate kernel size. The situation in this case becomes simpler. Because $H_f(f_x, f_y; z)$ is to be multiplied by the Fourier transform of the hologram $T_h(f_x, f_y; z_o)$, we can now take both the frequency sample spacings and bandwidth of $H_f(f_x, f_y; z)$ to be the same as those of $T_h(f_x, f_y; z_o)$. That is

$$\begin{aligned} \delta f_x &= \frac{f_{xmax}}{NX} = \frac{L_{ox} + L_{hx}}{2NX\lambda z} \\ \delta f_y &= \frac{f_{ymax}}{NY} = \frac{L_{oy} + L_{hy}}{2NY\lambda z} \\ N_{Kx} &= NX, \quad N_{Ky} = NY \end{aligned} \tag{3.18}$$

where NX and NY are the sample number of the hologram in x and y dimensions, and f_{xmax} and f_{ymax} are the maximum frequencies of the hologram in each frequency direction as in (3.6) and (3.5) respectively. The δf_x and δf_y are frequency sample spacings of $H_f(f_x, f_y; z)$ in each dimension, and N_{Kx} and N_{Ky} are the total sample numbers. Note that if the hologram is not optimally sampled, rather, it chooses the smallest values of X and Y, then according to section 3.1 $f_{xmax} = 1/(2X)$ and $f_{ymax} = 1/(2Y)$, it yields:

$$\begin{aligned} \delta f_x &= \frac{f_{xmax}}{NX} = \frac{1}{2L_{hx}} \\ \delta f_y &= \frac{f_{ymax}}{NY} = \frac{1}{2L_{hy}} \\ N_{Kx} &= NX, \quad N_{Ky} = NY \end{aligned} \tag{3.19}$$

In summary, the discretized frequency kernel can be represented as

$$\begin{aligned} H_f[l, k; z] &= \iint h_f(x, y; z) \exp\{j2\pi(l\delta f_x x + k\delta f_y y)dxdy \\ &= \exp\{j\pi\lambda z[(l\delta f_x)^2 + (k\delta f_y)^2]\} \end{aligned} \tag{3.20}$$

where $l = -N_{Kx}/2, ..., N_{Kx}/2 - 1$ and $k = -N_{Ky}/2, ..., N_{Ky}/2 - 1$.

The frequency domain kernel can also be computed by doing the FFT of $h_f[m, n; z]$. If we denote the FFT of $h_f[m, n; z]$ as $FH_f[l, k; z]$, it can be expressed as (note the origin has

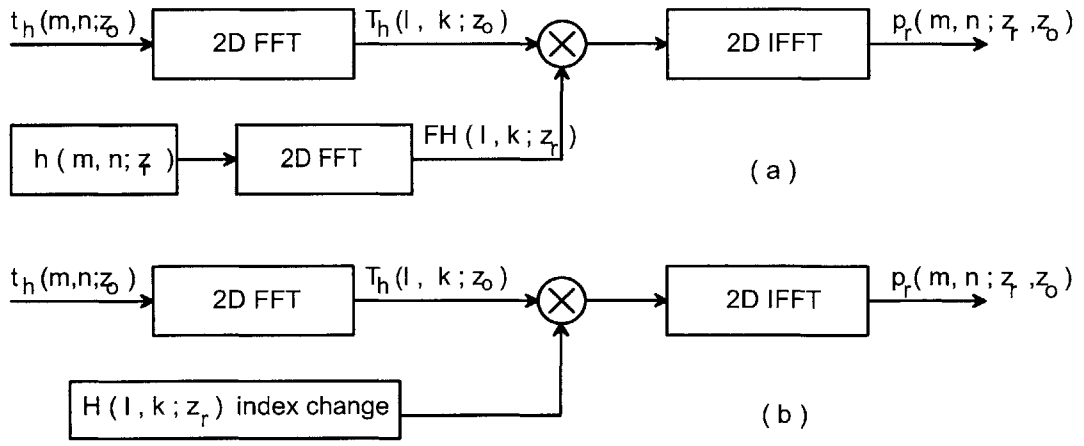

Figure 3-3: Algorithm structure; (a) start with spatial kernel construction; (b) construct frequency kernel directly.

been shifted):

$$\begin{aligned} FH_f[l,k;z] &= \sum_{i=0}^{NX-1}\sum_{j=0}^{NY-1} h_f[i-NX/2, j-NY/2; z]\exp(-j2\pi(\frac{li}{NX}+\frac{kj}{NY})) \\ &= \exp(-j\pi(l+k))\sum_{m=-NX/2}^{NX/2-1}\sum_{n=-NY/2}^{NY/2-1} h_f[m,n;z]\exp(-j2\pi(\frac{lm}{NX}+\frac{kn}{NY})) \end{aligned} \tag{3.21}$$

where $l = 0, ...NX-1$ and $k = 0, ...NY-1$. Note that $FH_f[l,k;z]$ is different from $H_f[l,k;z]$. $H_f[l,k;z]$ samples the Fourier transform of $h_f(x,y;z)$, while $FH_f[l,k;z]$ is the sampled discrete Fourier transform of $h_f[m,n;z]$. Also caution must be paid to the index difference between them when they are being exchanged.

3.2.2 Algorithm Structure

Once we have the discrete representation of both the hologram and the kernel, algorithms of computational implementations of optical reconstruction simply follow as multiplying the 2D FFT of the hologram with one of the frequency kernels, and then do the inverse FFT. Their structure is shown in Figure 3-3.

3.2.3 Behavior Of The Reconstruction Process: Focus Analysis

From chapter 2 we know that reconstruction leads to a virtual image and a real image. Reconstructed at the distance $z_r = -z_o$, the real image becomes an out-of-focus twin-image and the virtual image comes into focus. However, in most cases the value of z_o is an unknown parameter to be estimated. Consequently, reconstruction errors often originated from two different sources: the twin images of the objects whose virtual image are focused and the out-of-focus images of other objects.

In this section, we first describe qualitatively the behavior of the computational reconstruction by calculating the virtual image and real image profiles of a simple object geometry, for both in-focus and out of focus situations. The twin image effect will then be illustrated by separate discussions of each component. Then several focus measures will be proposed in attempt to quantitatively characterize the variation of the reconstructed images around the focus point. For simplicity, we use paraxial approximation and in some places we neglect the effect of finite hologram aperture in all the derivations. To further simplify the analysis, a simple object distribution function is chosen as:

$$t_o(x,y;z_o) = \begin{cases} 1 & \sqrt{x^2+y^2} \le a \\ 0 & \text{otherwise} \end{cases}$$

which is a 2D circular disc of radius a.

First let's treat the variations of the virtual image and twin-image separately, as the reconstruction distance passes by the focus distance. Numerical results of focus analysis with and without considering the twin-image effect will be presented.

Denoted as $u_1(x_r,y_r;z_r,z_o)$, the virtual image term can be expressed as (cf.equation (2.26)):

$$\begin{aligned} u_1(x_r,y_r;z_r,z_o) &= \beta A^2\{t_o(x,y;z_o) * * h_f(x,y;z_r+z_o)\} \mid_{x=x_r,y=y_r} \\ &= \beta A^2\{t_o(x,y;z_o) * * \frac{1}{j\lambda(z_r+z_o)} \exp(j\frac{\pi}{\lambda(z_o+z_r)})\} \mid_{x=x_r,y=y_r} \end{aligned} \tag{3.22}$$

Or in the frequency domain,

$$U_1(f_x,f_y;z_r,z_o) = T_o(f_x,f_y;z_o)H_f(f_x,f_y;z_r,z_o)$$

$$= T_o(f_x, f_y; z_o) \exp\left[-j\pi\lambda(z_r + z_o)(f_x^2 + f_y^2)\right] \tag{3.23}$$

When $z_r = -z_o$, since we have,

$$\begin{aligned} H_f(f_x, f_y; z_r, z_o) &= 1 \\ h_f(x, y; z_r; z_o) &= \delta(x)\delta(y) \end{aligned} \tag{3.24}$$

Therefore, from (3.22),

$$\begin{aligned} u_1(x_r, y_r; z_r = -z_o) &= \beta A^2 t_o(x, y; z_o) * *\delta(x)\delta(y) \\ &= \beta A^2 t_o(x_r, y_r; z_o) \end{aligned} \tag{3.25}$$

Hence the virtual image is exactly reconstructed up to a constant factor.

However, if $z_r = -z_o + \delta z$ and thus the image is misfocused by δz, then according to (3.22), the reconstructed virtual image becomes,

$$\begin{aligned} u_1(x_r, y_r; z_r, z_o) &= \beta A^2 \{t_o(x, y; z_o) * * h_f(x, y; \delta z)\} \mid_{x=x_r, y=y_r} \\ &= \beta A^2 \{t_o(x, y; z_o) * * \frac{1}{j\lambda\delta z} \exp(j\frac{\pi}{\lambda(\delta z)}(x^2 + y^2))\} \mid_{x=x_r, y=y_r} \end{aligned} \tag{3.26}$$

That is, the misfocused virtual image becomes a set of holographic fringes parameterized by δz, and is independent of z_o.

Computed according to equation (3.25) (3.26) and using the circular disc object, a plot of the focused virtual image is given in Figure 3-4(a), which is identical to the object. In Figure 3-4(b) the profile of the virtual image defocused by $\delta z =$ is shown after being enlarged. In the same figure the object profile is plotted in dashed lines. We can see that the defocused virtual image is seriously distorted especially around the edges.

The expression of twin-image can be derived in the same way hence is not repeated here. The twin-image associated with the focused virtual image has been computed separately and its profile is shown in Figure 3-5(a). We could see that the twin-image component fuly degenerates into a set of fringes and sharers no similarity to the original object shape which is plotted in dashed lines. The total reconstructed image, consisting of both the virtual image and twin-image, has been computed and its profile is plotted in Figure 3-5(b). The

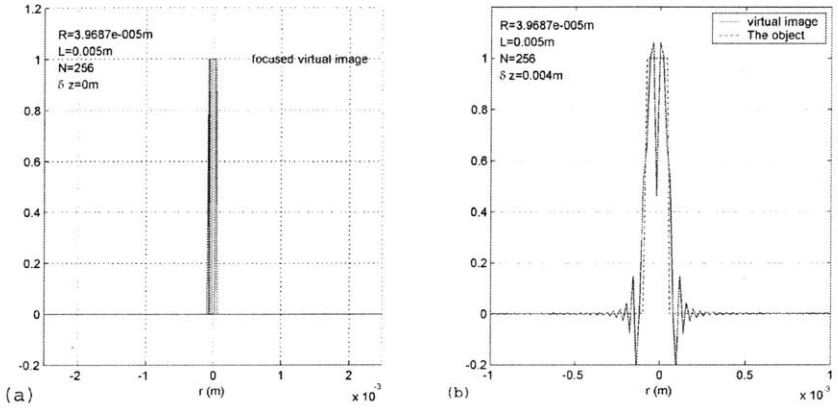

Figure 3-4: Profiles of the virtual image of a circular object
(a) In focus; (b) Defocused by δz.

degradation of the reconstructed image due to the twin-image can be explicitly seen in Figure 3-5(b), as compared to Figure 3-4(a).

The above discussion and plots illustrates how each reconstruction component varies around the focus point and how the twin-image effect comes into play. Nevertheless, this is not enough to precisely characterize the focus sensitivity of a reconstruction. A quantitative focus analysis would be preferable here which usually proceeds in terms of certain focus measures. A focus measure is often defined such that it is maximum at the best focused image and decreases as the misfocus increase. Then the curve of this focus measure verses the misfocus distance becomes a good characterization of this variation. Many focus measures have been suggested in areas of computer vision and microscopy, mainly for auto focus analysis. Which one to use really depends on the specific application. A good review of focus analysis and focus measures can be found in [22], where three focus measures have been proved to be theoretically sound, including the image energy, the energy of image gradient and the energy of Laplacian. All of these have been shown to be unimodal, monotonic focus measures. Focus analysis for holography application was also discussed by R. Yu et.al. [28]. Rather than one single measure, they proposed a set of focus measures each corresponding to a different wavelet decomposition level. The response of measures associated with lower levels is relative flat. As the levels increase, the focus measure could achieve a sharpest response at one of these levels. This multi-resolution approach is equivalent to defining a

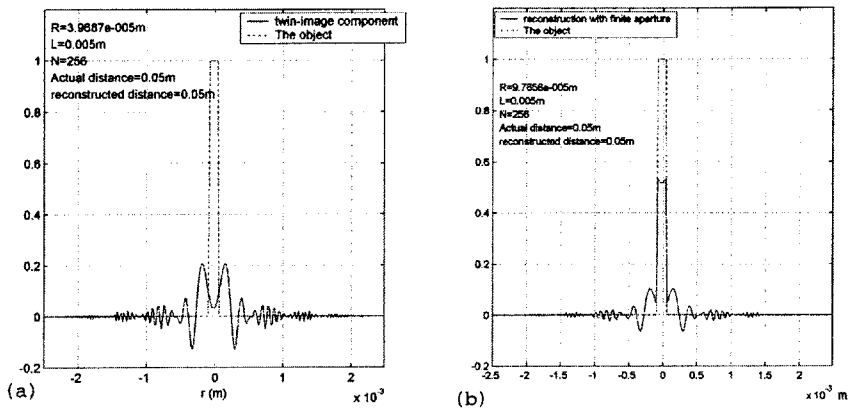

Figure 3-5: the twin-image effect
(a) The profile of the twin-image associated with the focused virtual image;
(b) The total reconstructed image associated with focused virtual image.

focus measure for each frequency band of the image.

Now for our particular example of a circular object, we could for analysis purposes define a focus measure that is related to the mean square error between the reconstructed image and the circular disc object. In this sense, reconstruction is equivalent to a matched filtering. The output is then the reconstruction. The mean square error is due to the misfocus. Assuming $\beta A^2 = 1$, The expression of MSE follows directly from (3.26), for virtual image only:

$$\begin{aligned} MSE &= \iint \{t_o(x,y;z_o) - u_1(x,y;z_r,z_o)\}^2 dxdy \\ &= \iint \mid T_o(f_x,f_y;z_o) \mid^2 \mid 1 - exp(-j\pi\lambda\delta z(f_x^2+f_y^2)) \mid^2 df_x df_y \\ &= \iint \mid T_o(f_x,f_y;z_o) \mid^2 2\{1 - cos\pi\lambda\delta z(f_x^2+f_y^2)\} df_x df_y \end{aligned} \tag{3.27}$$

The same calculation applies when the total image is concerned. We can see that for the virtual image only, the MSE consists of errors contributed from each frequency component. When reconstructed in the best focus, the virtual image is assumed to be the exact replica of the object. Thus the MSE becomes zero. Note that in equation (3.27) we have not considered the finite aperture effect, which, as we've discussed in chapter 2, will actually distort the virtual image even in the best focus. Thus the realistic value of MSE is not zero at the focus point. We then define the focus measure as the reciprocal of the MSE

error. After running the reconstruction through a sequence of distances around the focus point and calculating the MSE of the virtual image component and the total image at each distance, the curves of MSE-based focus measure are plotted in Figure 3-6. For these, the integration is over a finite aperture. It is in logarithmic scale after normalization. Figure 3-6(a) gives the numerical results of the MSE focus measure for virtual image only, and Figure 3-6(a) gives the result of the total field. In computation the following setup has been used : object radius $a \approx 40\mu m$, object distance $z = 50\ mm$, hologram aperture $L = 5\ mm$. The MSE focus measure of the virtual image only has a sharp response resembling a Delta function, with resolution less than $0.5\mu m$. This is understandable since the focused virtual image is almost identical to the object except the difference due to the finite aperture effect. However, the response of the MSE focus measure for the total field is much flatter, with resolution at the order of half mm. This can be attributed to the presence of the twin-image component.

In addition to the MSE focus measure, we have also computed the results using an energy concentration measure and a gradient peak measure. The corresponding results for both the virtual image and the total image field have been plotted in Figure 3-7 and Figure 3-8 respectively. The energy concentration measure is defined by the total reconstructed energy inside the nonzero object region:

$$EC = \iint |t_o(x,y;z_o)u_1(x,y;z_r,z_o)|^2 dxdy \tag{3.28}$$

As shown in Figure 3-7, for both the virtual image and total reconstructed image, the focus measure has a resolution at the order of 1 mm. Thus it is less sensitive to the twin-image interference. The gradient peak measure is defined as the maximum local gradient value of the reconstructed image, thus for virtual image only:

$$GP = \iint \sqrt{(\partial_x u_1(x,y;z_r,z_o))^2 + (\partial_y u_1(x,y;z_r,z_o))^2}\, dxdy \tag{3.29}$$

where the partial derivatives are local operators. This is based on the assumption that when reconstructed in focus, an image has the largest gradient distribution. As shown in Figure

3-8, the gradient peak measure is also insensitive to the twin-image. It has a resolution at the order of 0.4 mm which is better than the energy concentration measure.

In summary, we could say that the twin-image could impose serious effects on the reconstruction focus sensitivity, depending on the focus measure operators. This effect, however, is not so obvious when one looks at a reconstructed image contaminated by the twin-image. Also interestingly we can see that, different focus measures give variable responses in this particular case. The choice of an appropriate focus measure, as one could expect, would be closely related to the imaging nature of holography technique.

3.3 Reconstruction Improvement by Kernel Modification

Previous discussion has shown that the effect of the twin-image imposes a out-of-focus background on the virtual image, and thus prevents good reconstruction quality. Also the strong directly transmitted light usually overlaps the object region and reduces contrast. Off-axis holography effectively separates the virtual image, twin-image and transmitted light one from another in space. However, that also means more complicated optical alignment and larger sampling frequency requirement for computational reconstruction. Optical high-pass filtering was suggested being able to reduce the depth of focus and suppress the speckle noise in in-line holography [26]. Recently in HPIV field, H. Meng proposed a so called in-line recording off axis viewing method (IROV) [17]. This IROV technique records and reconstructs an in-line hologram in the normal way. However, the observation angle to the reconstructed field is tilted to an angle, i.e. off-axis viewing. It was said that this off-axis viewing method could reduce the speckle noise and improve the axial focus resolution since it avoids the directly transmitted light and only uses the side scattering components of the object recorded by the hologram.

In this section, we present two different types of operation which could improve the reconstructed image quality by attenuating the forward scattering component, and are closely related to the IROV approach. They correspond to revisions of the reconstruction kernel in spatial and frequency domains. For in-line holography, the object light incident on the hologram can be divided into forward scattering (low spatial frequency) and side scattering (high spatial frequency) components. Typically the fringes due to the forward scattering have better contrast than those due to the side scattering. However, there are

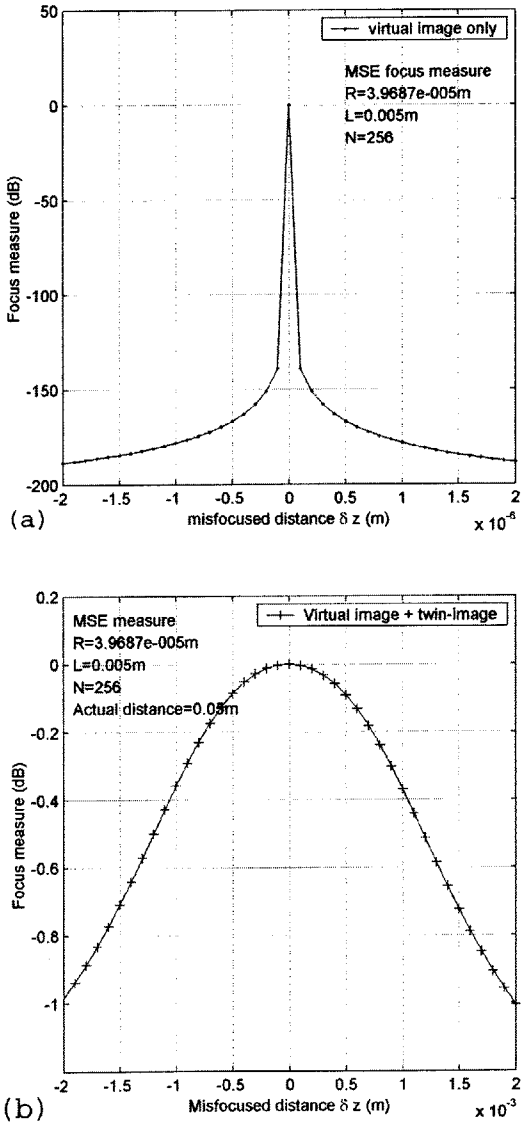

Figure 3-6: MSE focus measure verses misfocus distance
(a) Virtual image only; (b) Virtual image and twin-image

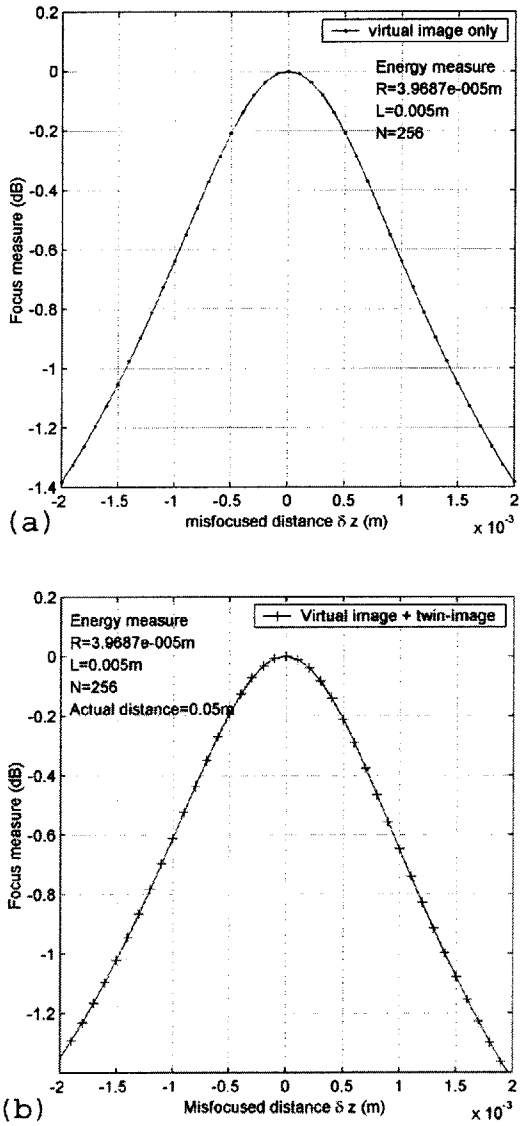

Figure 3-7: Energy focus measure verses misfocus distance
(a) Virtual image only; (b) Virtual image and twin-image

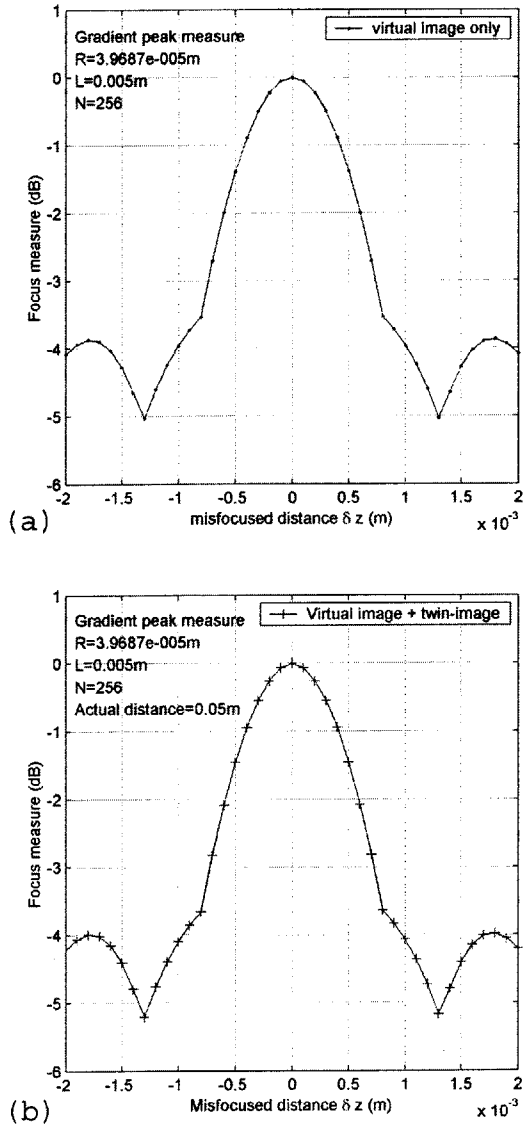

Figure 3-8: Gradient peak focus measure verses misfocus distance
(a) Virtual image only; (b) Virtual image and twin-image

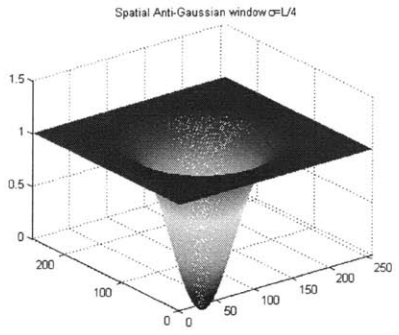

Figure 3-9: The spatial window function

two disadvantages when forward scattering is involved. First, The forward scattering light usually has a small angle with the reference light. Thus the focus resolution associated with these components are usually poor due to small numerical aperture. Secondly, the forward scattering field usually is contaminated by the strong transmitted light and halo. Therefore in this sense, the side scattering deserves more weight. This idea can be implemented both in spatial domain and frequency domains, by nullifying the low frequency portion of both the spatial and the frequency kernels.

We use Fresnel approximation for both the spatial and the frequency kernels in this section.

3.3.1 Kernel Windowing in the Spatial and Frequency domains

An anti-Gaussian window is imposed on the spatial kernel. The window is centered at zero and its shape can be adjusted by changing its "variance". Since the center portion of the kernel is associated with small spatial frequency (local spatial frequency), it is equivalent to high pass filtering. The window function can be represented by:

$$w(x,y) = 1 - \exp\frac{-(x^2+y^2)}{2\sigma^2} \tag{3.30}$$

Figure 3-14 shows the window function with $\sigma = L/4$, L is the kernel size in one dimension. Figure 3-13 shows the spatial kernel with and without windowing, together with their 2D profiles passing through the origin.

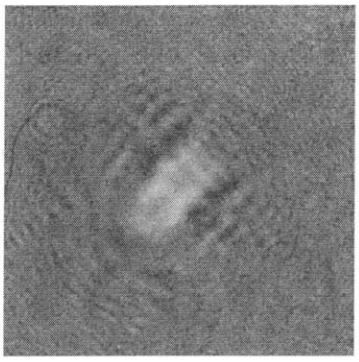

Figure 3-10: The hologram

A similar window function can be imposed to the frequency kernel. it is:

$$w(f_x, f_y) = 1 - \frac{1}{2\pi\sigma^2} \exp \frac{-(f_x^2 + f_y^2)}{2\sigma^2} \tag{3.31}$$

Figure 3-11 shows the frequency kernel with and without windowing, together with the 2D center profiles. The window function used in Figure 3-11 has $\sigma = F/4$, F is the sampling frequency in one dimension.

Both operations have been implemented on one real hologram which is shown in Figure 3-10. Its physical size is 8*8 mm. To find out the effect of windowing the frequency kernel, windows with different shapes have been tried and their results are shown in Figure 3-12. The corresponding results of spatial kernel windowing are shown in Figure 3-14. We can see that by selecting a suitable window shape, it is possible to suppress the bright background behind the object and thus improve the contrast near the object region. The reconstructed object image has a better intelligibility than the normal reconstruction.

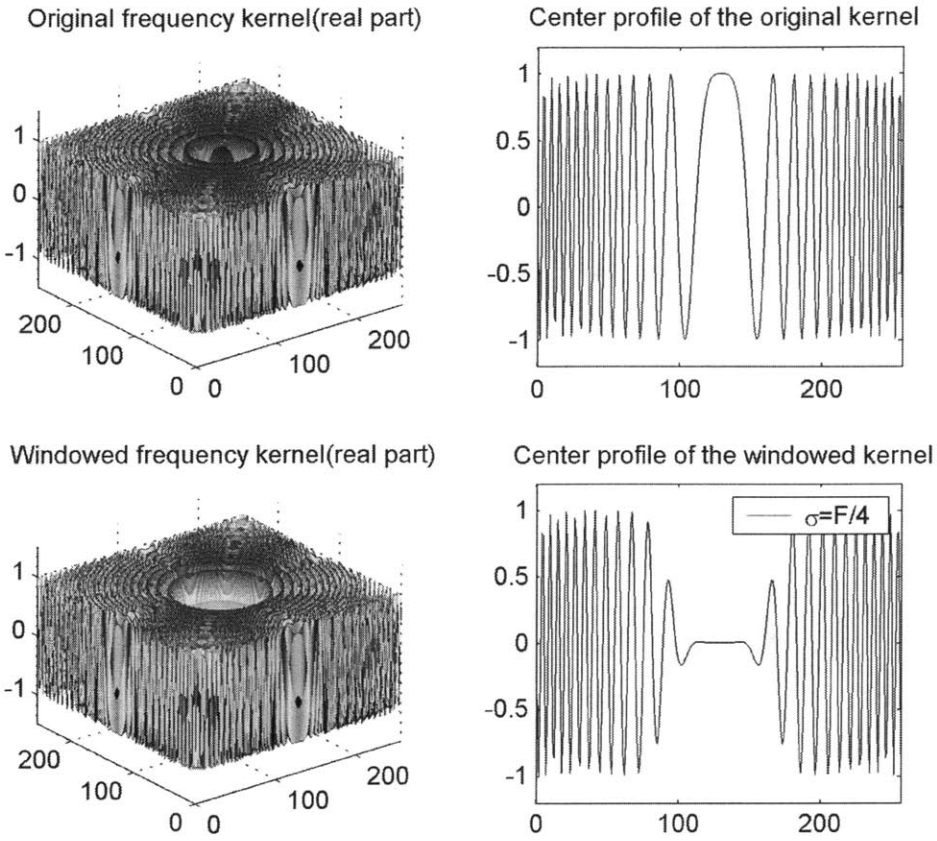

Figure 3-11: The frequency kernel with and without windowing (F is the sampling frequency)

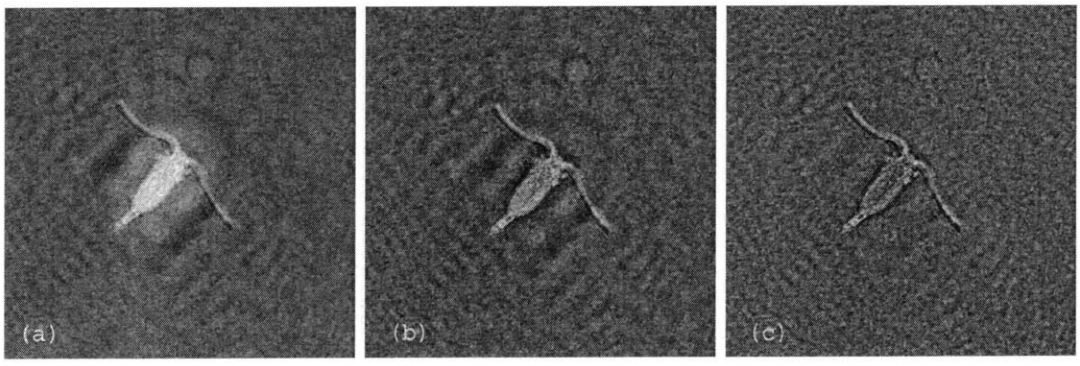

Figure 3-12: Reconstruction with different window shapes in frequency domain
(a) the normal reconstruction; (b) window shape $\sigma = F/64$;
(c) window shape $\sigma = F/32$ (F is the sampling frequency).

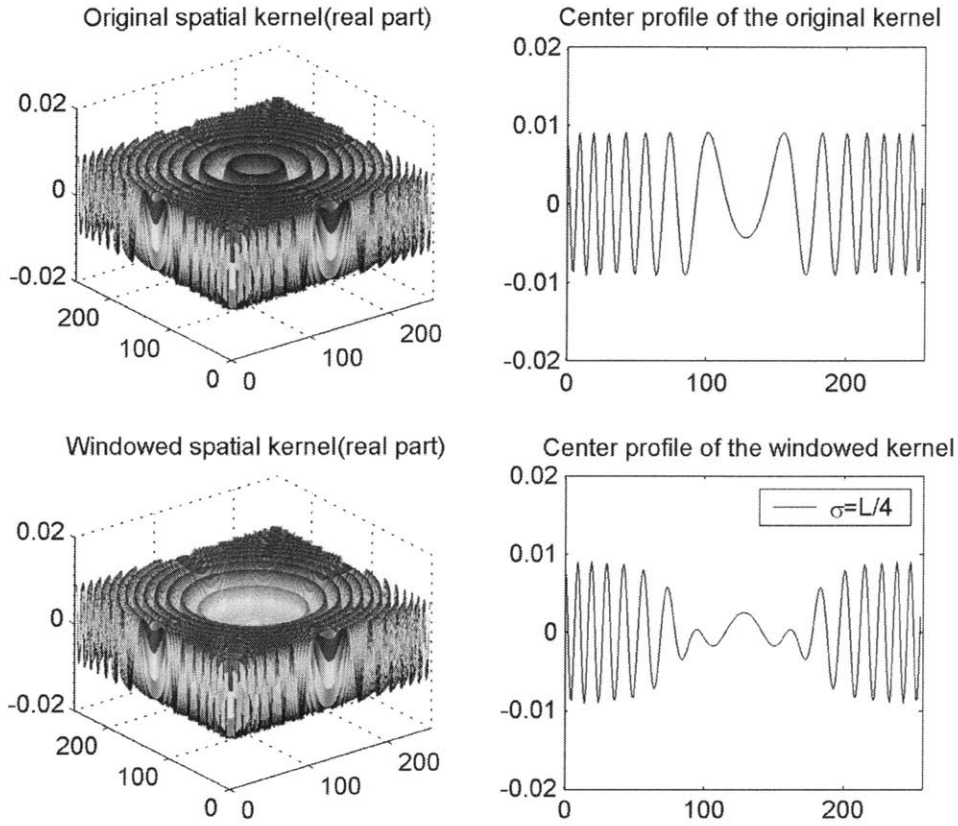

Figure 3-13: The spatial kernel with and without windowing (L is the kernel size in one dimension)

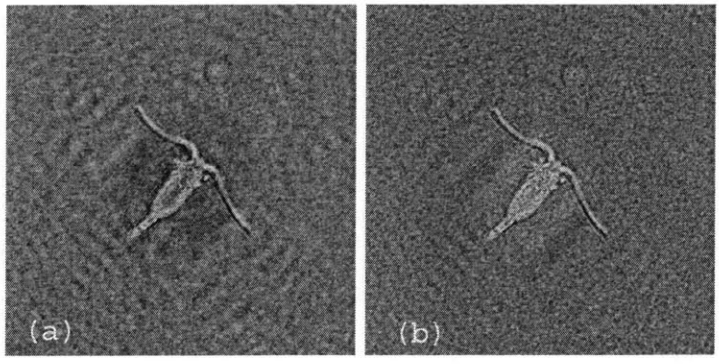

Figure 3-14: Reconstruction with different window shapes in spatial domain
(a) window shape $\sigma = L/32$;
(b) window shape $\sigma = L/8$ (L is the kernel size in one dimension).

Chapter 4

Digital Reconstruction System

So far we have discussed the mathematical formulation of holographic reconstruction in Chapter 2 and computational reconstruction algorithms in Chapter 3. In this Chapter, we will focus on system implementation issues. First, a conventional optical reconstruction system will be briefly reviewed. Then the DSP based computational reconstruction system will be introduced in detail, including its hardware and software structure.

Often, optical reconstruction results in real and virtual three dimensional images of the optical field that initially made the hologram appear in space to someone viewing it. Although these images are visually perceptible, they are not easily used scientifically. In scientific or industrial holography applications, such as HPIV (Holographic Particle Imaging Velocimetry, see section 5.1.1) and underwater hologrammetry (see section5.2.1), people are particularly interested in quantitative analysis of the hologram and its reconstructions. Hence it is desirable to digitize the optical field and record the digitized data for use in subsequent observation and processing. One example of this is measurements of plankton and distribution of their sizes done by Malkiel et al. (1999) [15]. Another example of the process is given by Zhang et al. (1997) [29] who measured the 3-dimensional turbulent fluid motion of the flow in a duct having a square cross section. The flow was "seeded" with particles which were imaged in a double exposure hologram to carry out 3-D holographic particle image velocimetry (HPIV). In both examples, digitizing and computer recording of the data was done by using a video camera, fitted with microscope optics, mounted on a computer-controlled stage that moved in three dimensions as the data was digitized and recorded. This is typical of conventional optical reconstruction. As technology evolved,

Table 4.1: Number of real multiplications involved in one reconstruction, without considering kernel construction

Operation	Size	number of real Multi.
2D FFT of hologram	6400*6400	$2 * 6400^2 log_2(6400)$
Matrix Multiplication	6400*6400	6400*6400*4
2D IFFT	6400*6400	$2 * 6400^2 log_2(6400)$

digital reconstruction systems emerged.

Conventional Optical Reconstruction System: This type of system positions digital imaging device at the far end. The system completes optical reconstruction first. The reconstruction is then scanned in three dimensions by a digital imaging device on a computer-controlled mechanical platform. Figure 4-1 gives an example of this type of system.

Computational Reconstruction System: Optical holograms are initially recorded and developed on analog media such as holographic film or glass-plates; then they are scanned for subsequent computational reconstruction.

Film-Free Digital Reconstruction System: In this type of system, holograms are directly imaged by CCD cameras instead of analog media. Then they are computationally reconstructed and processed accordingly. No film and chemistry needed.

These systems have been limited by several technological factors. First, CCD sensors with large aperture size and small pixel size have long been desired but have only recently become available at affordable prices. This prevents digital holography from being widely used. Currently CCD sensors with 4080*4080 pixel number and $9\mu m * 9\mu m$ resolution are available from Kodak. But it is too expensive for extensive usage. Secondly, the speed and memory required of the computational platform is substantial. For large scale applications with required resolution of a few microns, the amount of image data and amount of computation become large. Table4.1 gives the number of real multiplications and the amount of data that are involved in reconstruction of one image from a 2 inch by 2 inch hologram which is sampled at 3200 dpi, i.e. 7.9 microns.

So without considering the kernel construction, the total number of real multiplication of one reconstruction is $2.24 * 10^9$. Assume the hologram's depth of volume to be 200mm,

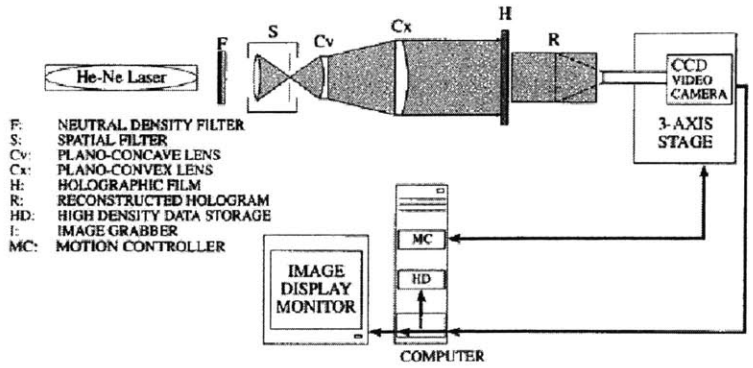

Figure 4-1: Conventional Optical Reconstruction System(courtesy of Dr. E. Malkiel)

and the depth separation between consecutive images to be 0.1mm. Then the full representation of object volume requires 2000 reconstructed image planes. The total number of real multiplications goes up to $4.48 * 10^{12}$.

In addition, if each image is in 8bits gray scale mode, then a single image size is $6400 * 6400 = 4.096\ GegaBytes$. The space needed to hold all the reconstructions is $2000 * 6400 * 6400 = 81.92\ GegaBytes$.

All of these are only for reconstructions of a single hologram. Hence real time computational reconstruction is far beyond the capability of most current computational platforms. However, recent advances in DSP technology has reduced or eliminated the computational barrier. As a first step towards film-free real-time holographic system, we have a DSP based computational reconstruction system which will be presented in section 4.2.

4.1 Review: Optical Reconstruction and 3D Scanning

As mentioned above, conventional reconstruction system tends to complete the whole reconstruction optically and then use a mechanically controlled digital camera system to scan the three dimensional optical field. Most of todays's holographic replay systems are still in this form, such as in the Laboratory for Experimental Fluid Dynamics at Johns Hopkins Univ. and European HOLOMAR system.

Figure 4-1 illustrates the setup of a system which is currently being used in the Labora-

tory for Experimental Fluid Dynamics, Johns Hopkins University. The system consists of reference laser source, lens system and CCD video camera with microscope optics mounted on a three degree platform which is controlled by a desktop computer. The controlling software guides the camera through sections of reconstructed volume and allow the user to stop and fine tune the camera position when an object comes into focus. Each section may need several scans due to the limited aperture size of the video camera. The user can select the interested region which usually contains focused objects and store the image on the computer.

This system has several advantages. First, the reconstructed volume has very high in-plane resolution and depth resolution due to the optical reconstruction. Second, image storage size can be effectively reduced by selectively picking up regions where objects come into best focus. An intensity measurement method has been suggested in [15] to determine focus position of objects. Third, object classification analysis can be implemented after the image is captured.

However, two major facts about this system limit its usage and actually motivated digital reconstruction. One is that the scan speed is very slow. According to [15], usually it takes 100-400 hours to finish one scan of a single hologram. The other issue is stability. The quality of optical reconstruction relies strictly on the platform stability. During field operations such as on a ship at sea, such a long time scan is very difficult.

In addition, there is a potential detriment in the system depth resolution. Though the minimum resolvable depth distance between two optically reconstructed objects can be of the order of 10 microns depending on the wavelength, object size, hologram aperture and recording distance [2], image sequence recorded by video camera does not attain such a high depth resolution due to the limited depth of field of the camera and lens system that are being used, which often can be at the order of 100 microns.

4.2 Digital Reconstruction System

A feasible digital reconstruction system must have capacity in four aspects, including: sufficient digitization resolution, high computation speed, enough intermediate data buffer and large disk space for reconstructed images. Our computational reconstruction system has been constructed such that all these requirements can be satisfyingly met. Its main

components are selected to address each of the above four issues respectively:

Imacon FlexTight Photo Film Scanner 3200 dpi CCD drum scanner

Alacron FastImage DSP Board consists of 4 Philips TriMedia TM1300 chips

Alacron FastMem Board 512MB PMC Memory card

A high-end PC with 100GB disk space

The stand alone film scanner is capable to do scans at a resolution finer than 8 μm which is the typical particle size in HPIV and suitable for some underwater hologrammetry. The PMC memory card has been attached to the DSP board which is connected to the host PC via PCI bus. The work flow to reconstruct one optical hologram is like the following: First, the hologram is scanned at selected resolution and sent to the host PC where it is stored as a digital image. The user then can specify the reconstruction depth sequence, wavelength and particular reconstruction method to be adopted. Then the host PC transfers the hologram image data to DSP system together with the user's input, and initiates the computation on the DSP board. During the period of reconstruction, each image reconstructed at one distance will be swapped to host PC disk space. The PMC memory space allows the reconstruction image size up to 4096*4096, or 8192*8192 with the help of the host PC's memory space.

However, the most dramatic improvement happens to the reconstruction speed, which is attributed to the DSP engine. While optical reconstruction typically takes 100-400 hours to reconstruct from one hologram, the computational system takes 70 seconds to generate one 4096*4096 reconstruction image and less than 4 hours to get 200 images at different distances. Strategy like coarse search followed by fine-tune focus search can be adopted to find out accurately the distance where each object comes in to focus. Consequently the total reconstruction time of one hologram is reduced by the factor of 10, i.e. from several hundred hours to a few hours.

4.2.1 The DSP Reconstruction Engine

The memory-enhanced DSP system is the computation engine of the reconstruction system. It is configured with four TM1300 trimedia DSP chips which work simultaneously during the reconstruction. The DSP board communicates with the host PC via PCI bus through which

Figure 4-2: DSP Based Digital Reconstruction System

the hologram image data and user specified parameters are imported and the reconstructed images are sent back. All the computations are completed in the DSP system except for image display.

The program running in the DSP engine is essentially the algorithms that have been presented in Chapter 3. Actually under the author's help, programmers at Alacron based the DSP C code directly on an earlier PC program written by the author. In terms of reconstruction of one image plane, the computation mainly involves the 2D forward FFT of the hologram image, the kernel generation in wavenumber space, the multiplication of the kernel with the Fourier transform of the hologram, then 2D inverse transform of the products and the image gray scale normalization. To generate a series of reconstruction images, the Fourier transform of the original hologram needs only to be calculated once and then can be reused in sequent reconstructions.

For DSP implementation, however, due to the parallel processing structure, the whole processing task has to be partitioned and distributed to each processor node with a suitable data path layout. Processes running on all nodes have to be synchronized. A flow chart of the program is given in Figure 4-3. As seen in Figure 4-3, the whole program consists of 6 phases. Phase 1 and Phase 2 compute the 2D forward FFT of the hologram image to be reconstructed. All the 2D FFT's and inverse FFT's in the program are first computed in rows and then in columns. Phase 2 also constructs the kernel function. If the exact kernel is selected, it is first generated in the spatial domain and then transformed into the wavenumber space via 2D FFT. If the paraxial approximation kernel is to be used, it is then directly constructed in the wavenumber space since the analytical form is known, and then is multiplied with the FFT of the hologram image. Phase 3 finishes the exact kernel construction and multiplies it with the hologram in wavenumber domain. In Phase 4, an

intermediate reconstruction is generated by 2D inverse FFT of the product and its dynamic range is measured. After normalization, Phase 5 generates the final reconstructed image which is in 8 bits or 16 bits gray scale format. Phase 3 to Phase 6 make up a loop which is useful for reconstructions at a sequence of distances. At each distance the kernel function is re calculated in Phase 6. However, the 2D FFT of the hologram needs only to be computed once. At the beginning of each Phase, all the nodes are synchronized by the host PC. In Phase 6, the host PC reads back the reconstructed image and updates the parameters for the next loop including the new distance value.

The speedup of reconstruction time due to the DSP engine is obvious as can be seen in Figure 4-4. Note that as the image size increases, the reduction becomes more substantial and is more than a factor of 10 at 8192*8192 as compared to PC implementations.

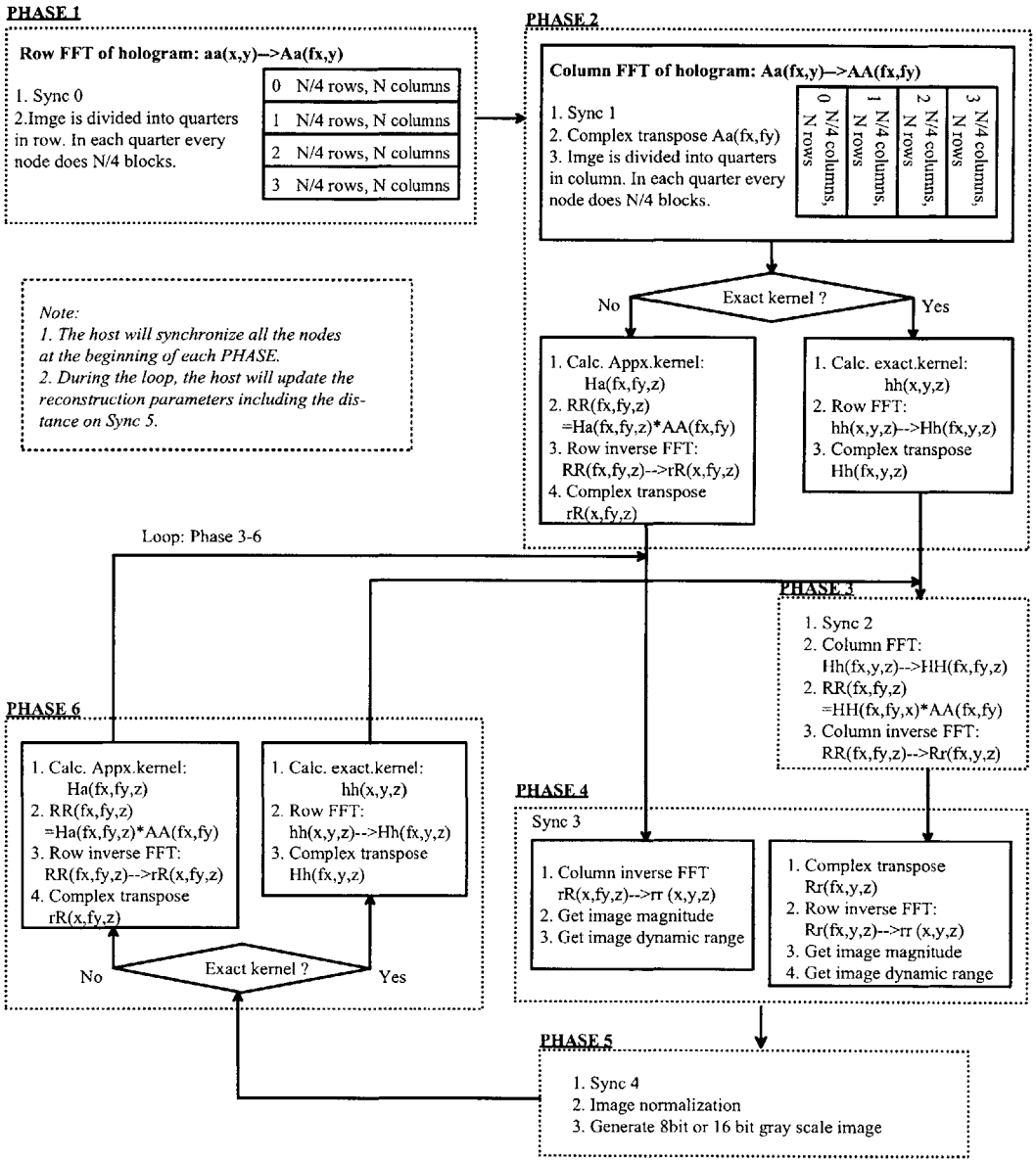

Figure 4-3: DSP program flowchart

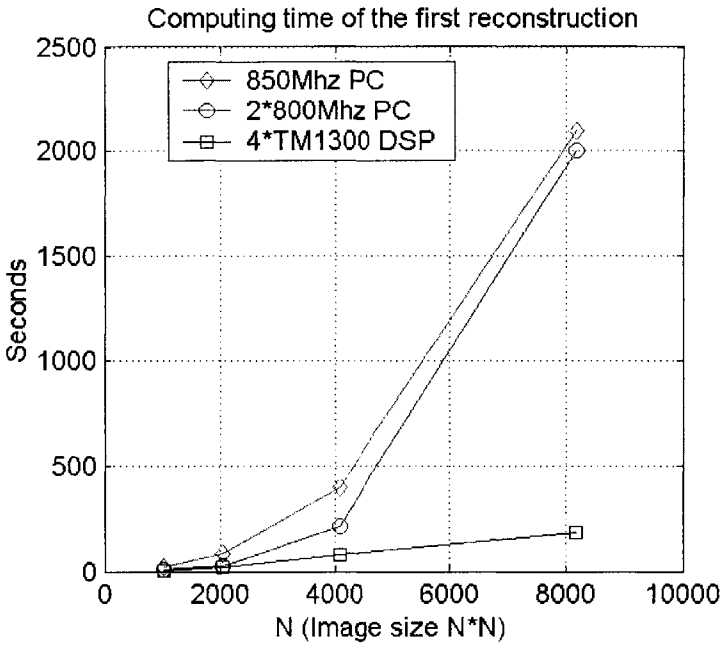

Figure 4-4: Single reconstruction time on different computing platforms

Chapter 5

Simulation and Experimental Results

In this chapter, we will present results in two kinds of applications: HPIV(Holography Particle Imaging Velocimetry) and underwater hologrammetry.

For HPIV, synthesized particle holograms of a water volume are generated and then reconstructed by applying algorithms presented in chapter 3. Scenarios with different particle densities have been tried. In fact near the focus, images reconstructed from synthesized particle holograms are illustrative examples to show the behavior of the reconstruction process.

For underwater hologrammetry, images of marine micro-organisms computed from holograms will be shown that are made with Johns Hopkins University submersible holocamera (c.f. Katz, at al.1999) by the JHU staff. In most of these examples their optical reconstructions will be shown as well so comparisons can be made. Finally, a procedure will be presented that has been devised to determine the micro-organism population densities contained in a volume from its hologram.

The content of this chapter is organized as following: In section 5.1, after a brief review of HPIV technology, procedures to synthesize particle holograms are introduced. Then reconstructions of the synthesized hologram are shown associated with different particle density. In section 5.2, example images computed from plankton holograms will be shown. The influence of scan resolution on the final image quality will be illustrated with results computed from a same optical hologram scanned at 3200 dpi and 8000 dpi respectively. Spatial windowing and filtering operations have been applied on some examples and the

results will be shown. In section 5.2.3, an object counting technique is described which substantially reduces the computation time by using a summation kernel.

5.1 Reconstruction of Synthesized Particle Holograms

5.1.1 Introduction to HPIV

Particle Image Velocimetry (PIV) which measures two components of velocity in a 2D plane, has become a state-of-the-art experimental technique for fluid velocity measurement over the past decade [1]. However, in cases when instantaneous full field high-resolution velocity measurements are demanded which are resolvable both in space and time , the information capacity requirement is far beyond the PIV and photography-based techniques. Holography comes into play as an inherently 3D imaging process. Holographic PIV (HPIV) records the instantaneous 3D information of a large quantity of tracer particles in a fluid volume on a hologram and then reconstructs the particle images optically. By finding the 3D displacements of the particles in the image volume between two exposures separated by a very short time lapse, an instantaneous volumetric 3D velocity field is retrieved.

The key points of optical HPIV are the reconstruction quality and effective means to handle the huge quantities of data, namely to get the volumetric velocity vector field. Some of these issues have been discussed and reported with optical implementations. In [18], Meng et. al. discussed the speckle noise problem in in-line particle holography. In [17], the same author also presented an in-line recording and off-axis viewing technique to reduce the twin-image effect of in-line particle holograms. Zhang et al. [29] reported a hybrid HPIV system which measures turbulent flow in a square duct. An optically advanced HPIV system has recently been developed by Pu and Meng (2000) [21] and applied to flow measurements in both air and water. Each of these complex systems has a unique optical setup which usually lacks user-friendliness. To fully explore the capability of a HPIV system, it is necessary to reduce the system complexity and increase the measuring speed. With the help of digital recording technology, the computational reconstruction system together with the algorithms presented in previous chapters therefore provide a premising solution due to its high-speed reconstruction as well as its simple and stable setup.

5.1.2 Synthesized Particle Holograms and Reconstructions

For this work, synthesized particle holograms of water volume in a square duct are generated. Two alternative source models are assumed by the synthesis program.

The first model is particle-based. It takes each particle inside the volume as a circular disk with programmable in-plane radius and zero thickness along the optical axis, at a specified distance from the hologram plane. Then the hologram is a superposition of contributions from all particles inside the volume. The contribution of each particle is the interference pattern between the light scattered by that particle and the transmitted light. The angular spectrum of the scattered field can be computed by using Bessel Fourier transform due to the circular symmetry of the assumed particle geometry. Then it is multiplied by the diffraction kernels presented in Chapter 2. Finally the interference pattern of that particle is computed as the magnitude of the inverse FFT of that product. Here again we assume that particle density is sparse such that the incident light can pass through, and that the strength of the scattered light is much weaker than the transmitted light. Thus that the mutual-interference between particles and multi-scattering are negligible as well as the nonlinear terms.

The second model, however, is plane-based. The volume is discretized along optical axis into a sequence of planes parallel to the hologram and equally separated byΔz. The nth plane is at distance z_n from the hologram. Then particles falling into $(z_n - \Delta z/2,\ z_n + \Delta z/2]$are mapped onto the nth plane. The hologram then is the superposition of contributions from all planes. The contribution of each plane can be calculated following the same procedure of computational reconstruction, which involves twice 2D FFT's and convolution with the diffraction kernel. The same assumptions hold here as for the first model.

The synthesis program is based on an earlier GUI version developed by C. P.Earls [3]. The water volume is modeled to be of a square duct shape which has a square cross section of $8mm * 8mm$ and thickness of 12.8mm. This is actually smaller than most HPIV system setups, but it is sufficient to demonstrate effects and results. For all the results to be shown below, the plane based model has been used. The volume is divided into $n = 256$ planes with interval $\Delta z = 0.05mm$. Each plane is discretized into 2048*2048 pixels with every pixel counting for 3.9*3.9 square microns.

Figure 5-1 illustrates the three-dimensional view of the sample volume. The program

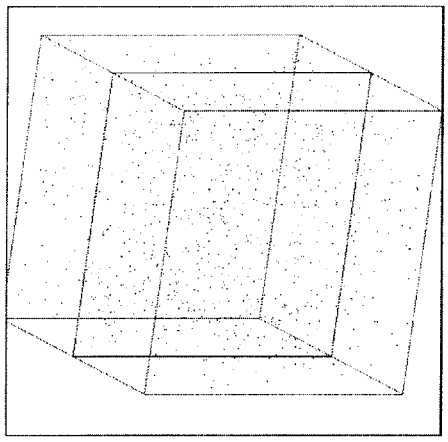

Figure 5-1: The 3D view of the volume

is designed such that particles can be "seeded" randomly into the volume with uniform distribution statistics and programmable density. It also allows the user to input particles one by one with specified coordinates.

To visualize the reconstruction process, a small set of signal particles are deliberately arranged into a pattern of character "H" which has the size around 0.54mm by 0.54mm. The signal particles are surrounded by noise particles randomly scattered inside the whole volume. All signal particles have a radius of 10.4 micron and are on a same plane at 100mm from the hologram. All noise particles have a raduius of 4 microns and their locations are random. The simulation is done with three different scenarios each with noise seeding density of 0, 2.4 and 5 particles/mm^3. Computational reconstructions are then calculated for cases near the focus plane at 100mm.

Figure 5-2(a) gives a plane view of the sample volume whose seeding density is 5 particles/mm^3. The signal particle pattern is hardly visible due to the dense noise particle distribution. Hence in figure 5-2(b), the center portion of (a) is shown after enlarged by a factor of 4, in which the signal particles are clearly visible in a pattern of "H". Note that the plane view actually projects the whole volume onto one plane parallel to the hologram. Particles shown in the plane view are not all in one plane.

Figure 5-3 shows the synthesized hologram of a noiseless volume and the associated reconstruction images near the focus plane. For better observation, only the center portion of each reconstructed image is shown after enlargement. Figure 5-3(a) is the hologram image.

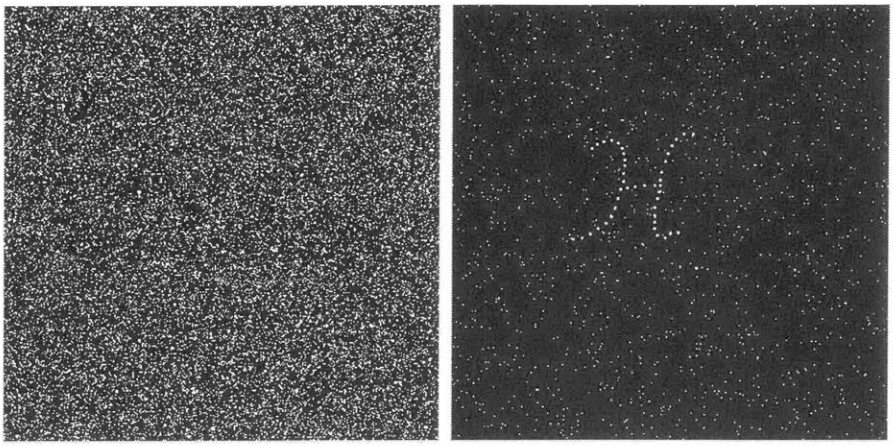

Figure 5-2: a. The plane view of the sample volume.
b. The central portion of (a) after enlarged by a factor of 4.

The aliasing in the synthesized hologram is due to the frequency domain implementation of the circular convolution without zero-padding. However, this aliasing does not introduce visible distortion in reconstructed images as given in Figure 5-3(b)-(f), each of which are computed at$z$ = 99, 99.5, 100, 100.5 and 101 mm respectively. The image sequence is symmetric with regard to z=100mm where the particles are in the best focus. We can see how the image of each particle moves out of focus and becomes a set of circular fringes as the distance moves away from the focus distance.

The twin image effect is not substantial as we can observe from these images.

Now in addition to the set of signal particles, we "inject" noise particles into the volume with density of 5 particles per mm^3. Figure 5-4 shows the synthesized hologram of the noisy volume and the associated reconstruction images after enlargement. The hologram is shown after contrast and brightness increase both of 30. However, it still has worse fringe contrast than the noiseless case due to the existence of noise particles. Consequently, the quality of reconstructed images Figure 5-4 (b)-(f) is inferior to that of Figure 5-3 (b)-(f).

Another example is shown in Figure 5-5. The noise particle density is 2.4 particles per mm^3. The hologram in this case has a better fringe contrast as compared to Figure 5-4 (a), which however is worse than the fringe contrast of Figure 5-3(a).

The simulation described above does not perfectly match the realistic HPIV setup which often has a larger geometry and suffers the influence of speckle noise. However, the appli-

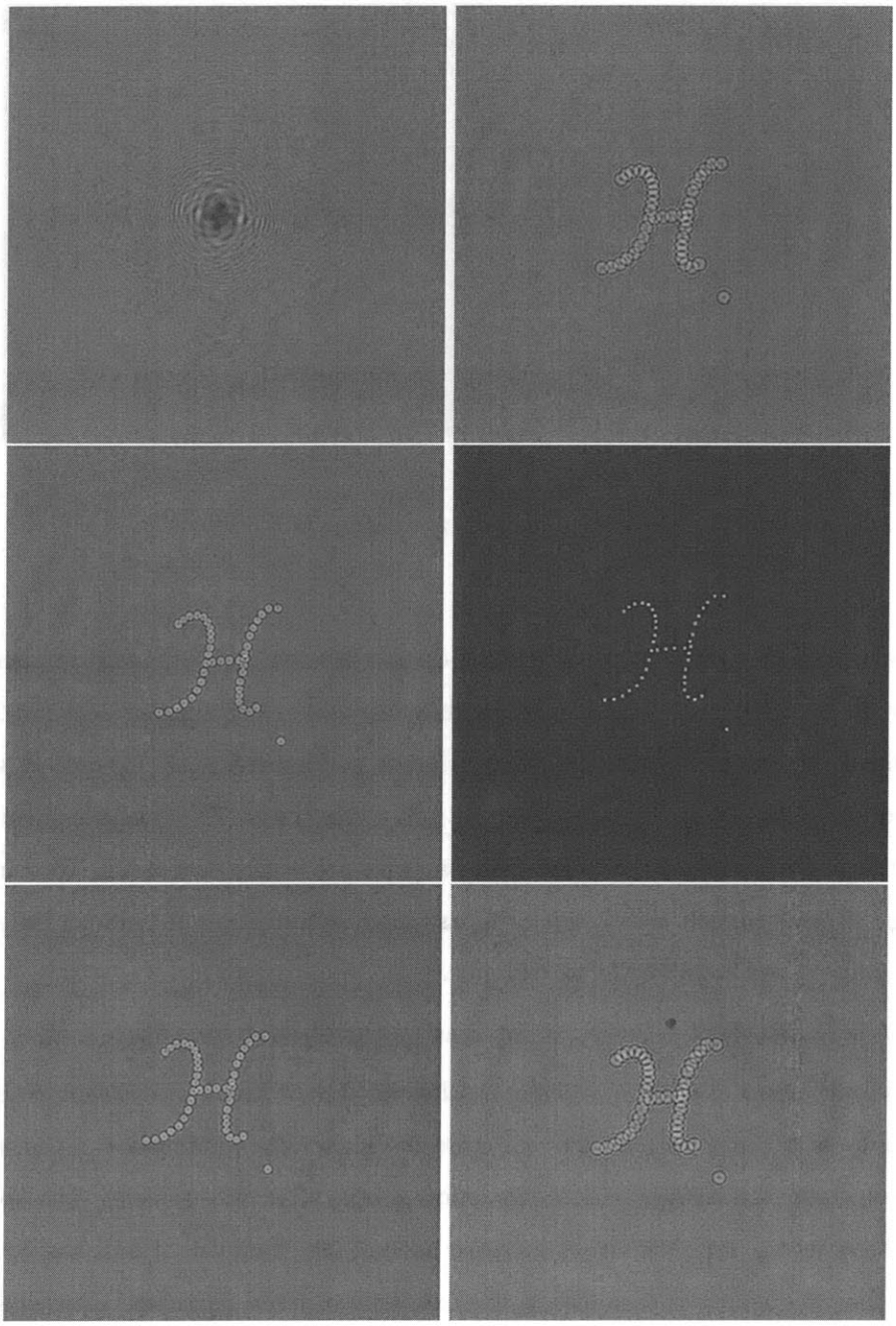

Figure 5-3: The hologram of volume without noise and its reconstructions
a. The hologram.
b. Reconstructed image plane at 99mm from the hologram.
c. Reconstructed image plane at 99.5mm from the hologram.
d. Reconstructed image plane at 100mm from the hologram.
e. Reconstructed image plane at 100.5mm from the hologram.
f. Reconstructed image plane at 101mm from the hologram.

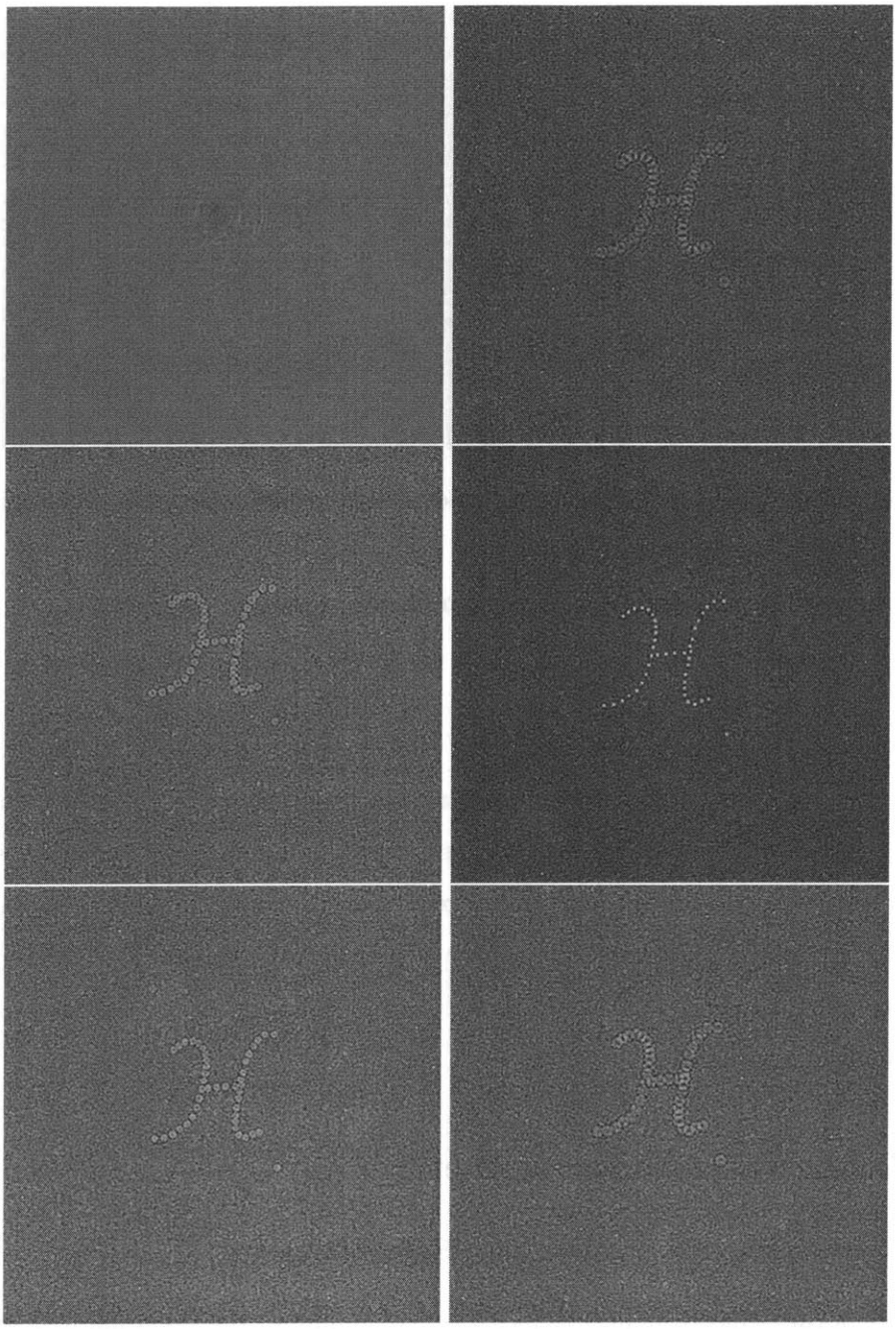

Figure 5-4: The hologram of volume with noise density of 5 particles per mm^3 and its reconstructions
a. The hologram.
b. Reconstructed image plane at 99mm from the hologram.
c. Reconstructed image plane at 99.5mm from the hologram.
d. Reconstructed image plane at 100mm from the hologram.
e. Reconstructed image plane at 100.5mm from the hologram.
f. Reconstructed image plane at 101mm from the hologram.

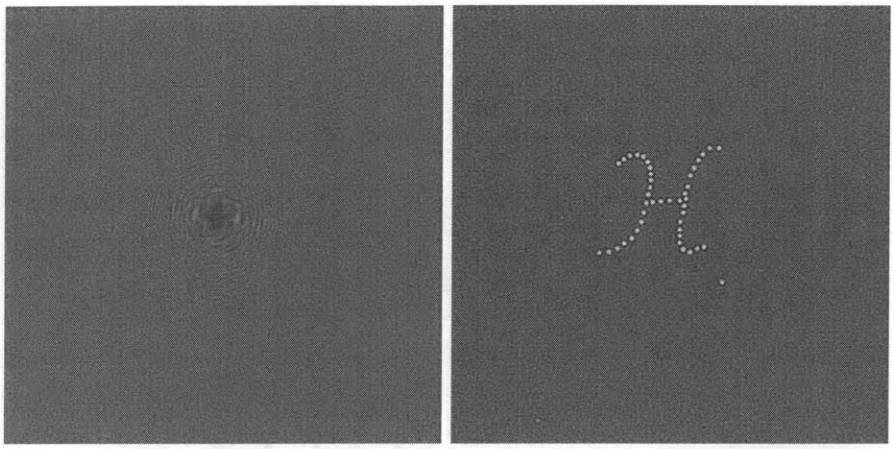

Figure 5-5: The hologram of volume with noise density of 2.4 particles per mm^3 and its reconstructions
a. The hologram.
b. Reconstructed image plane at 100mm from the hologram.

cability of computational reconstruction can still be clearly inferred from these examples. Furthermore, combined with a digital imaging device, film-free digital HPIV system is a promising path for future HPIV development, which in addition to providing simplified system setup, will also provide high-speed processing capability including reconstruction and volumetric velocity retrieval. The barrier is, however, the availability of high resolution CCD sensors with sufficient pixel count and a computation platform that can handle the huge amount of computation.

5.2 Underwater Hologrammetry

5.2.1 Introduction to Underwater Hologrammetry

Knowledge of the behavior of marine biological communities, their relation with each other and with the particles with which they interact is very important for many reasons including sustainable fishery resources. Also it is widely accepted that the earth's climate is intimately connected with and dependent on processes occurring in the seas and oceans. The underlying control forces of these process are nearly always biotic involving the growth, death, and sedimentation of the diverse array of plants and animals including plankton in the sea. The accuracy of the conceptual and mathematical modeling of these bio-geochemical

processes suffers at present from inadequate data. This is largely due to the lack of appropriate techniques for making the necessary observations. Appropriate models cannot be developed unless reliable and accurate, non-invasive techniques are available to study the spatial relationships of the biotic and abiotic particles of the ocean. The overriding benefit of underwater hologrammetry is that it allows the non-invasive, in-situ, volume recording of living organisms and inanimate particles in their natural environment. This has hitherto been difficult or impossible to obtain.

Currently underwater hologrammtry records in-situ holograms of aquatic systems using pulsed laser and subsequently replays in air the images of reconstruction. The images, so produced, are life size, fully three-dimensional and located in real image space in front of the observer. Such images can be directly interrogated to extract information from the hologram on the particles, including their morphology, identity and positions.

In the last five years, systems of different dimensions and diverse optical setup have been developed and reported. One example is from the HOLOMAR project in Europe. The HOLOMAR project involves research groups from five different countries. Their holocamera is a hybrid system with both in-line and off-axis configurations. The system and some of its applications in the study of plankton distributions have been published with satisfying experimental results [5] [12]. Watson et al. [27] discussed issues about optimization of reconstructed images such as aberration reduction and index compensation. Another holocamera system has been developed by the Laboratory for Experimental Fluid Dynamics at Johns Hopkins University. It is an in-line system and probably the first for this applications. Experimental results of measurements of plankton distribution in ocean have been reported [15].

All these systems use analog media for recording and optical replay for subsequent measurements. The main problem, as already mentioned in chapter 3, is that the process is laborous, time-intensive and sensitive to the environment. During experimental deployment, it is strategically important to get immediate feedback from the first several deployments, such that the subsequent deployment plan can be adjusted if necessary to achieve the best results without waste of time and resources. However, it is very difficult to get good optical replay results on board during experiments, especially in rough seas.

The computational reconstruction method as discussed in previous chapters, is essentially a beneficial substitution to optical replay to avoid all these problems. Since all re-

constructions are to be done computationally, the platform can be made less sensitive to environmental factors such as vibrations thn an optical system. The high-speed reconstruction is capable to quickly generate images as feedback for deployment. The M.I.T. computational system was used in a sea experiment in the Gulf of Maine in 2000 Summer, together with the JHU holocamera. During the field trip, the computational system generated images of plankton within minutes of development of the holocamera film.

5.2.2 Computational Reconstruction of Plankton Holograms

The current computational reconstruction basically takes digitalized optical holograms and generates planar images corresponding to cross sections of the imaged water volume. Subsequent measurements proceed by interrogating images one by one, to find out the identity of each particle by its individual morphology and to calculate the density of interested species contained in the volume using certain object counting techniques.

Before going any further in presenting some examples, the problem of index compensation is addressed first. In underwater holography, the hologram is recorded in water. If images are replayed optically in air, the difference in refractive index between the recording medium (water) and the replay medium (air) will introduce aberrations, astigmatism and change the field curvature and magnification. The concept of "index compensation" was introduced by Watson et al. [27] to overcome the effect of aberration. Its principle is that when the laser beam enters the water during recording it undergoes an effective wavelength change due to the refractive index of water. The effective wavelength λ_w can be expressed as:

$$\lambda_w = \frac{\lambda_a}{n_w}$$

where $n_w$is the refractive index of water and λ_a is the wavelength of the laser in air. So if the hologram is replayed in air with a wavelength equal to λ_w, the geometry of reconstructed image will be the same as that of the real volume.

Taking this "index compensation" into consideration, the computational reconstruction needs to set the assumed reference laser wavelength parameter to be λ_w. This is certainly much easier than trying to find a real laser of this wavelength in optical reconstruction. In terms of its effect on the reconstruction image, the consequence of using a wrong wavelength, is simply a distance shift in the paraxial approximation approach since the wavelength is

solely coupled to the distance by λz. In the case of using exact kernel without paraxial approximation, the consequence is not simple.

Figure 5-6 shows optical and computational reconstructions of a *calanus* copepod zooplankton. This particular specimen is a large copepod, about 3 mm long. The reconstructions are done at a distance of 124.5 mm from the hologram. An optical reconstruction as well as computational reconstructions using the exact kernel and the paraxial approximation are shown. Also, a reconstruction using the twin image reduction algorithm of Onural and Scott (1987) is shown. This is an iterative algorithm based on the paraxial approximation with each iteration removing more of the twin image, but introducing additional "noise". Therefore, there is an optimum number of iterations. We have used four iterations, a number recommended by Onural and Scott, in making this image. The focus of various parts of the image varies because the mid-plane of the copepod does not lie exactly in the reconstruction plane. Only a small part, about 4 mm x 4 mm of the reconstruction is shown. The digitized portion of the hologram has cross sectional dimensions of 32.5 mm x 32.5 mm and the optical reconstruction is of similar size. The digitizer resolution was 126 pixels per mm. The Fourier transforms in the computational reconstruction process are done using 4096 x 4096 points.

The four images are very much alike, but upon close inspection small differences can be observed. The hair-like constructions along the antennae of the copepod are more observable in the optical reconstruction with a resolution of about 3 microns than in the computational reconstruction with a resolution of about 8 microns. The holographic film resolution is better than 3 microns, but the optical system used to make the optical reconstructions limits the resolution to 3 microns. The scanner used to digitize the hologram for the computational resolution has a resolution of 126 pixels per mm with an equivalent resolution of 8 microns. Also, the computational reconstruction using the exact kernel is slightly "sharper" than the reconstruction whose kernel is based on the paraxial approximation. The figure with the twin image reduction (d) is nearly identical to the figure using the paraxial approximation, but containing the twin image (c) because of the previously mentioned small influence of the twin image on small objects.

Figure 5-7 shows an optical reconstruction and a computational reconstruction using the paraxial approximation of a diatom in the Chesapeake Bay. The diatom is about 2.3 mm long with a diameter of about 0.09 mm. A small section of the hologram was digitized

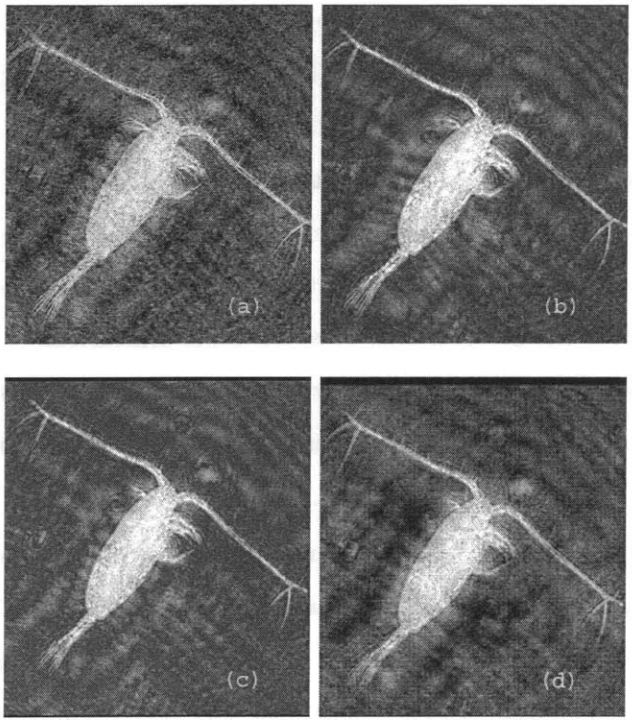

Figure 5-6: Image Reconstructions from a hologram of a large copepod.
a. Optical reconstruction (courtesy of Dr. E. Malkiel).
b. Computational reconstruction using the exact kernel.
c. Computational reconstruction using the paraxial approximation.
d. Computational reconstruction with twin image reduction.

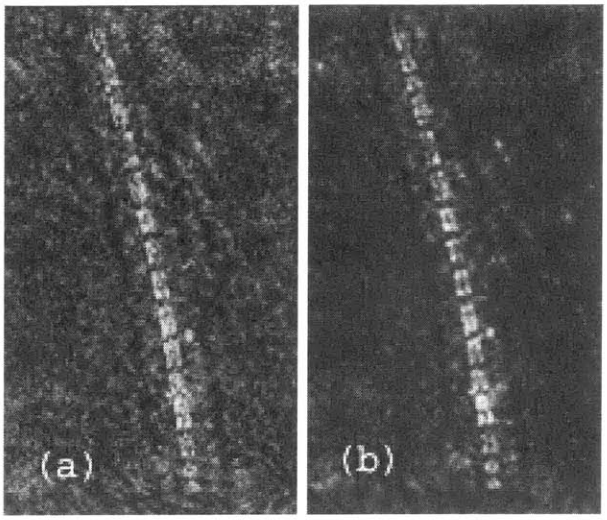

Figure 5-7: Image reconstructions from a hologram of a linear diatom.
a. Optical reconstruction (courtesy of Dr. J. Zhang).
b. Computational reconstruction using the paraxial approximation.

with a resolution of 4.9 microns. The number of pixels was 640 wide x 484 high and this was padded with zeros so a 1024 x 1024 point FFT could be used for the computational reconstruction. Both the optical and the computational reconstructions show the individual cells of the diatom very clearly. Only relevant portions of the images are shown.

Figure 5-8 shows optical and computational reconstructions of a helical diatom having a length of 0.6 mm along the helical axis and a strand diameter of about 0.04 mm. The exact kernel was used for the computational reconstructions. Two different computational reconstructions were done. One was with a 32.5 mm x 32.5 mm section of the hologram digitized at 126 pixels per mm with the FFT done using 4096 x 4096 points. The other was a 24.8 mm x 24.8 mm section of the hologram digitized at 315 pixels per mm and with the FFT done using 8192 x 8192 points. With the small strand diameter, the resolution based on 126 pixels per mm with an equivalent resolution of 8 microns is inferior to the optical reconstruction having a resolution of about 3 microns. However, when the hologram is digitized at 315 pixels per mm, corresponding to a resolution of 3.2 microns, the computational reconstruction is nearly as good as the optical reconstruction. The individual cells of the diatom are clearly observable. Again, only relevant portions of the reconstructed images are shown. The portions chosen are ones having the individual organisms described. This

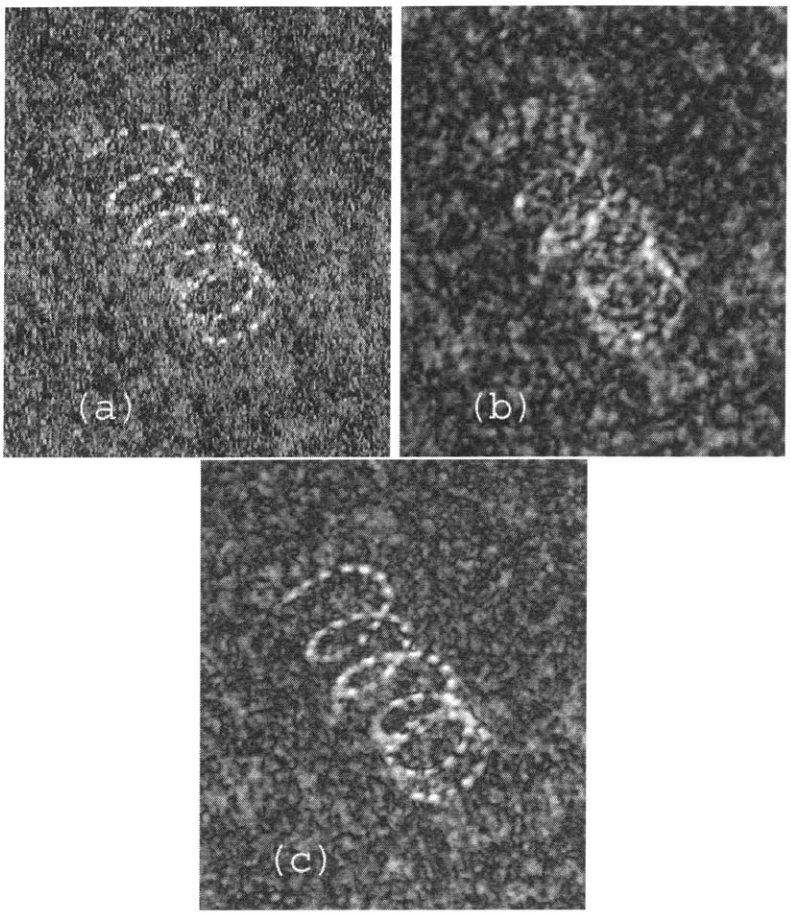

Figure 5-8: Image Reconstructions from a hologram of a helical diatom.
a. Optical reconstruction (courtesy of Dr. E. Malkiel).
b. Computational reconstruction with the hologram digitized at 126 pixels per mm.
c. Computational reconstruction with the hologram digitized at 315 pixels per mm.

example shows the improvement provided by better resolution when the number of pixels on a feature of interest is small. A resolution of 8 microns is equivalent to five pixels in the 40 micron size of the individual cells of the diatom, whereas the 3.2 micron digitizing resolution corresponds to about 13 pixels in the 40 micron size. These numbers are relative measures only since it is the hologram, not the image, that is digitized for the computational reconstruction.

Figure 5-9 shows optical and computational reconstructions of a plane through a copepod whose center plane lies at an angle to the reconstruction plane. The length of the copepod is about 0.7 mm. The resolution for digitizing the hologram was 315 pixels/mm. and

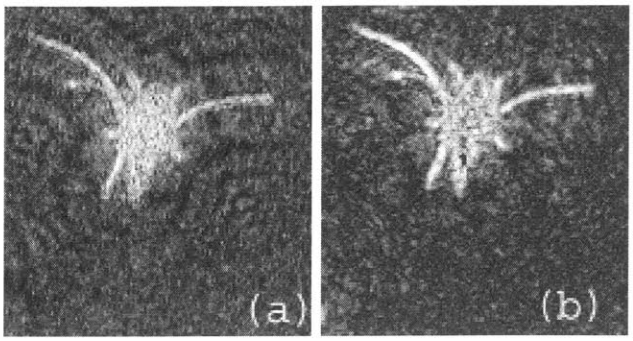

Figure 5-9: Image reconstructions of a copepod whose center plane is at an angle to the reconstruction plane.
a. Optical reconstruction b: Computational reconstruction with the hologram digitized at 315 pixels per mm.

8192 x 8192 pixels were processed. Since different parts of the copepod are at slightly different distances from the hologram, the focus varies slightly over the image. The optical and computational reconstructions are very similar which demonstrates the validity of the computational method at the shown resolution for objects of this size.

5.2.3 Object Counting Techniques

The JHU submersible Holocamera makes in-line holograms of volumes with a diameter of 50 mm and 250 mm long. Drs. J. Katz and E. Malkiel and their colleagues at JHU have devised a concept for determining the number of small objects in the imaged volume. Allowing for some fuzziness of the image, which is permissible for this application, the depth of focus of reconstructed images exceeds 1 mm. The concept is to make all the reconstructed planes parallel to the hologram, spaced about 2 mm apart and add up the intensities of all the planes. Each micro-organism will be close to in-focus in one or more planes and produce an out-of-focus image in other planes. The sum of all the planes will have a bright region in the vicinity of each micro-organism against a background of lower intensity. If the population density is small enough for overlaps of their images to be rare, counting the bright spots gives a close approximation to the number of objects in the imaged volume. The principal difficulty in doing this optically is the effort in making 124 plane image reconstructions for 2 mm spacing of the imaged planes.

Two approaches can be used to computationally generate the sum of the reconstructions.

One is to generate the image at each value of z and to add all the images together. The other is to take advantage of the fact that the image reconstruction is a linear signal processing operation. Making the sum of the imaged planes can be accomplished in the same way as making a single plane reconstruction, but replacing the reconstruction kernel, in either the spatial or the Fourier domains, by the sum of the kernels for each plane. Using the exact kernels, one kernel for each distance must be computed numerically and all the kernels then added together. The kernels are the sum of the ordinary kernels for values of z equal to $\delta z, 2\,\delta z, ...M\,\delta z$. Thus, for the exact kernel, the kernels in the spatial and frequency domains n_s and N_s respectively are given by:

$$n_s(x,y,z) = \sum_{m=1}^{M} n(x,y,m\,\delta z) \tag{5.1}$$

$$N_s(x,y,z) = \sum_{m=1}^{M} N(f_x,f_y,m\,\delta z) \tag{5.2}$$

Although this is reasonably computationally efficient, a considerable gain in efficiency is possible with the paraxial approximation for which the sum of the kernels can be determined analytically with that final expression used in the computations. The kernel in the Fourier domain for the paraxial approximation is given in equation 2.14. We need to generate the sum of the kernels for $z = \delta z, \delta z ... M\delta z$. This takes the form of the partial sum of a power series and is given by:

$$N_{ps} = \frac{i2\pi}{k} \sum_{m=1}^{M} \exp\left[-j\frac{2\pi^2 m\,\delta z}{k}(f_x^2+f_y^2)\right] = \begin{cases} \frac{j2\pi}{k}\frac{Q-Q^{M+1}}{1-Q} & Q \neq 1 \\ \frac{j2\pi}{k}M & Q = 1 \end{cases} \tag{5.3}$$

where:

$$Q = \exp\left[-j\frac{2\pi^2\delta z}{k}(f_x^2+f_y)^2\right] \tag{5.4}$$

Equations 5.1 through 5.3 are equivalent to using the same brightness and contrast scales for each of the images in the sum. On the other hand, if adding up the individual images is used, the intensity can be scaled for each image in the sum such that the darkest point is the lowest digital value used in the process (0) and the brightest point is the highest digital value used (255). Figure 5-10 shows these two cases with the individual images made using the paraxial kernel. For these images $\delta z = 2$ mm. The two images are nearly

identical, but they both have a contrast and image signatures that are insufficient for reliable computational counting of the objects. A number of steps are required before the objects can be reliably counted. The first step in preparing a summation image for computer-based object counting is removal of the low frequency intensity variation throughout the image. To do this, a bicubic equation in the two dimensions is formed with minimum mean square error from the summed image intensity and this is subtracted from the summed image, point-by-point. Then the image is scaled to cover the full dynamic range of 0 to 255. Figure 5-11 shows the result of this low frequency removal and scaling from figure 5-10b. Then a thresholding operation is applied to generate a binary image in which all points with intensities greater than a prescribed thresholding value have binary value 1 (shown white) and all other points have binary value 0 (shown black). Choice of the thresholding value is subjective and is best determined by visual examination of the results for several values. Thus the counting process is somewhat approximate. Figure 5-12 shows the results of applying a thresholding value of 0.38 of the maximum intensity to Figure 5-11.

Depending on just what one wishes to count, this process can give errors due to very small "noise spots". If the small extra white spots are thought to be "noise" and not something to be counted, they can be removed by *erosion*. In this process, any point is given a value of 1 only if it and a prescribed number of surrounding points have the value of 1. For example, Figure 5-13 is the result of applying a 12 pt x 12 pt erosion range to Figure 5-12. The corresponding physical erosion size is 0.1 mm x 0.1 mm. Some very small spots, which may be smaller than the printer can show, have been removed. It is possible for the erosion operation to break one object into two or more parts. The remaining objects can be enlarged for easier counting and there is a chance of "repairing" broken objects by a *dilation* process. In this process each point with a value of 1 is surrounded by a prescribed number of points in each direction having value 1.

Figure 5-14 shows the result of dilating Figure 5-13 by a box of size 20 x 20 pixels at each point. This corresponds to a physical size of 0.16 mm x 0.16 mm. It makes the remaining small objects easier to detect. The imaged volume of 2640 cm^3 is found to contain 24 objects with dimensions greater than 0.1 mm. The processes of erosion, dilation and counting were done using the functions in the MATLAB Image Processing Toolbox (1997). [16]

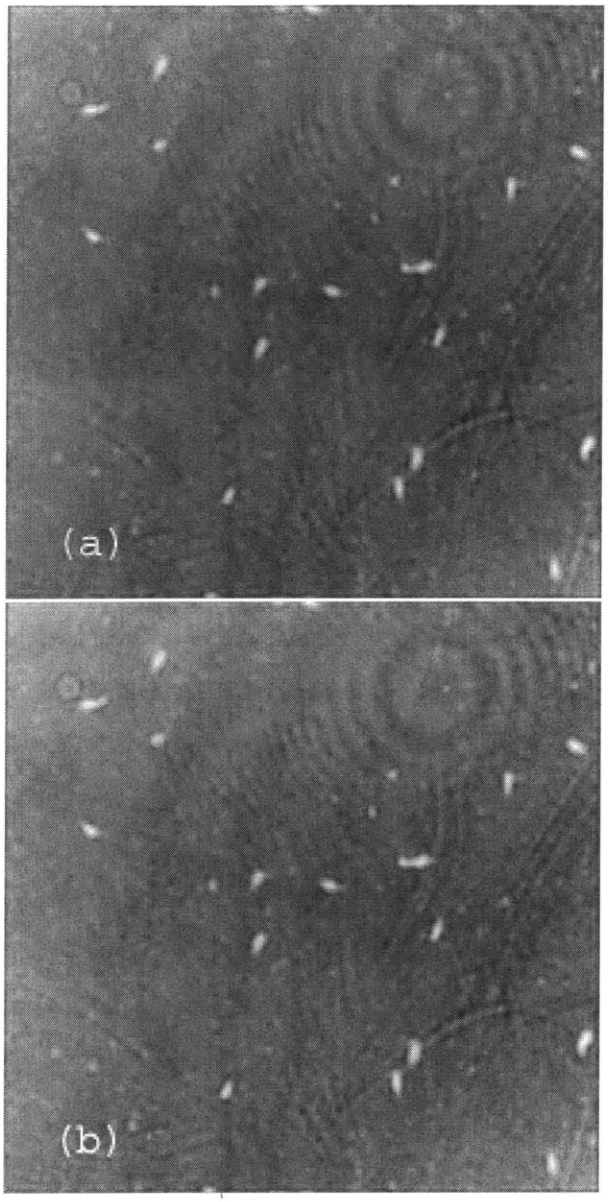

Figure 5-10: Summation of parallel planes through a holographic image, spaced 2 mm apart.
a. Each image has intensity scaled to full range.
b. Analytic sum for N_{ps} used.

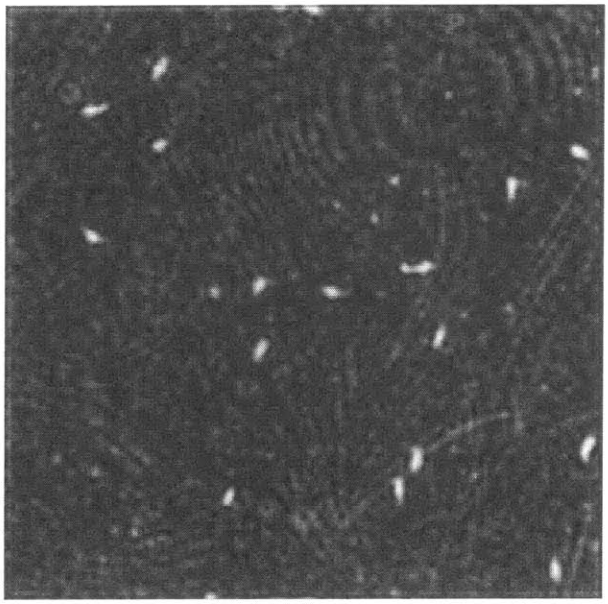

Figure 5-11: The image of Figure 5-10b after low frequency removal

Figure 5-12: Result of applying an 0.38 threshold level to Figure 5-11

Figure 5-13: Results of an 0.1 mm x 0.1 mm erosion process applied to Figure 5-12

Figure 5-14: The image of Figure 5-13 after dilation of each point to a square of dimensions 0.16 mm x 0.16 mm

Chapter 6

Conclusion and Future Directions

In summary, computational methods of reconstructing images from optical holograms have been developed and analyzed in this thesis. First a general mathematical formulation is derived for both in-line holography and off-axis holography. Issues involved in digital implementations have been discussed in detail including optimal sampling of hologram and discrete kernel reconstruction. The reconstruction behavior is described by considering the virtual image and twin-image individually. A number of focus measures have been proposed for focus analysis of the reconstructed image. Improvement of reconstructed image quality has been achieved by incorporating modified kernels. Secondly, a digital reconstruction system using DSP technology has been constructed. It speeds up the reconstruction time by an order of 10 compared to a PC platform and an order of 100 compared to optical reconstruction. Finally, both synthesized HPIV holograms and marine plankton holograms have been reconstructed and results are given showing that the computational methods and algorithms, together with the digital reconstruction system, are combined into a successful system that is capable of doing high-speed, high fidelity holographic reconstructions. Therefore our primal goal is achieved.

Problems remain, however, for future efforts.

First, with in-line holograms, there is still no sound solution to separating the focused image from its twin-image and images of other out-of-focus objects. Non-traditional filtering methods or novel transform domain approaches can possibly provide a way towards the answer.

Secondly, effective focus measures need to be developed and evaluated for auto-focus

analysis. It should be developed according to the physical natural of the holographic reconstruction. And according to author's own imagination, it is different from the focus definition of photography.

Third, computational reconstruction of off-axis holograms has not been explicitly discussed in this thesis, though its formulation is given. The effect of sampling in this case will not be the same in two dimensions.

Finally, from system point of view, efficient algorithms and powerful computation machines are in demand to bring computational reconstruction to real-time. For many applications, a film-free digital holographic system would be much more effective, efficient, robust and user friendly. This however can be achieved only when technology development loose the current limits of digital imaging devices.

Bibliography

[1] A.J. Adrain. Particle-imaging techniques for experiemntal fluid mechanics. *Annu. Rev. Fluid Mech. 23*, pages 261–304, 1991.

[2] R.A. Belz and F.M. Shofner. Characteristics and measurements of an aperture-limited in-line hologram image. *Applied Optics*, 11(10):2215–2222, 1972.

[3] Craig P. Earls. Holographic particle image velocimetry: Computational simulation and reconstruction. Master's thesis, Department of Ocean Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, 1999.

[4] N. L. Emmett and U. Juris. Reconstructed wavefronts and communication theory. *Journal of Optical Society of America*, 52(10):1123–1130, 1962.

[5] E. Foster and J. Watson. Holography for underwater inspection and measurement: an overview of current work. *Optics and Laser Technology*, 29(1):17–23, 1997.

[6] D. Gabor. A new microscopic principle. *Nature*, (4098), 1948.

[7] D. Gabor. Microscopy by reconstructed wave-fronts. *Proc. Roy. Soc.*, 197:454–487, 1949.

[8] D. Gabor. Microscopy by reconstructed wave-fronts:ii. *The Proceedings of The Physical Society (B)*, 64(378B):449–471, 1951.

[9] R. W. Gerchberg and W. O. Saxton. A practical algorithm for the determination of phase from image and diffraction plane pictures. *OPTIK*, 35(2):237–246, 1972.

[10] J. R. Goodman. *Introduction to Fourier Optics.* McGraw-Hill, New York, second edition, 1996.

[11] F. Gori. Fresnel transform, and sampling theorem. *Optics Communications*, 39(5):293–297, 1981.

[12] P.R. et al Hobson. A prelimimary study of the distribution of plankton using holograммetry. *Optics and Laser Technology*, 29(1):25–33, 1997.

[13] H. C. Van De Hulst. *Light scattering by small particles.* John Weley and Sons, New York, 1957.

[14] G. Liu and P.D. Scott. Phase retrieval and twin image elimination for in-line fresnel holograms. *Journal of the Optical Society of America A.*, 4(1):159–165, 1987.

[15] E. Malkiel, O. Alquaddoomi, and J. Katz. Measurements of plankton distribution in the ocean using submersible holography. *Measurement Science and Technology*, 10:1142–1152, 1999.

[16] The Math Works, Inc., 3 Apple Hill Drive, Natick, MA 01760-2098. *Image Processing Toolbox For Use with Matlab*, 1997.

[17] H. Meng. In-line recording and off-axis viewing (IROV) technique for holographic particle image velocimetry. *Applied Optics*, 34:1827–1840, 1995.

[18] H. Meng and W. L. et al Anderson. Intrinsic speckle noise in in-line particle holography. *Journal of the Optical Society of America A*, 10(9):2046–2058, 1993.

[19] L. Onural and P.D. Scott. Digital decoding of in-line holograms. *Optical Engineering*, 26(11):1124–1132, 1987.

[20] Ravikanth Srinivasa Pappu. Minimum information holograms. Master's thesis, Program in Media Arts and Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, 1995.

[21] Ye Pu and H. Meng. An advanced off-axis holographic particle image velocimetry (HPIV) system. *Experiments in Fluids*, 29:184–197, 2000.

[22] M. Subbarao, T. Choi, and A. Nikzad. Focuing techniques. *Optical Engineering*, 32(11):2824–2836, 1993.

[23] G. Tricoles. Computer generated holograms: a historical review. *Applied optics*, 26:4351–4360, 1987.

[24] A. VanderLugt. Optimum sampling of fresnel transform. *Applied Optics*, 29(23):3352–3361, 1990.

[25] Chandra S. Vikram. *Selected papers on holographic particle diagnostics.* SPIE optical engineering Press, 1990.

[26] C.S. Vikram. *Particle Field Holography.* Cambridge University Press, Cambridge, England, 1992.

[27] J. et al. Watson. Optimisation of holographic real images for subsea hologrammetry. *Proc. SPIE*, 2577:200–207, 1995.

[28] R. Yu, A.R. Allen, and J Watson. Focus measure for speckle noisey images based on wavelet multiresolution analysis. In *Wavelet Applications in Signal Prcessing V*, San Diego, 1997. SPIE.

[29] J. Zhang, B. Tao, and J. Katz. Turbulent flow measurements in a square duct with hybrid hologrphic PIV. *Experiments in Fluids*, 23(5):373–3819, 1997.